

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM protein - protein search, using sw model

Run on: April 21, 2005, 04:17:43 ; Search time 219.5 Seconds
(without alignments)
1185.131 Million cell updates/sec

Title: US-09-016-159D-5
Perfect score: 2715
Sequence: 1 MDHLGASLWPGVGSLLCLLA.....YENSLPAAEPLPVSIVACS 508

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2715	100.0	508	1 EPOR HUMAN	P19235 homo sapien
2	2283.5	84.1	509	1 EPOR_PIG	Q9myz9 sus scrofa
3	2174.5	80.1	507	1 EPOR_MOUSE	P14753 mus musculus
4	2152.5	79.3	507	1 EPOR RAT	Q07303 rattus norv
5	1670.5	61.5	387	2 Q95N13	Q95n13 ovis aries
6	1645	60.6	418	2 Q95N14	Q95n14 ovis aries
7	1058	39.0	316	2 Q35545	Q35545 rattus norv
8	1016	37.4	229	2 Q27950	Q27950 bos indicus
9	1011	37.2	229	2 Q28206	Q28206 bos taurus
10	529.5	19.5	529	2 Q6UAP7	Q6uap7 tetraodon n
11	251	9.2	522	1 IL9R HUMAN	Q01113 homo sapien
12	231.5	8.5	630	2 Q61Y8	Q61y8 gallus gall
13	224	8.3	566	2 Q925F5	Q925f5 mus musculus
14	220	8.1	625	1 PRLR_PIG	Q6jta8 sus scrofa
15	216.5	8.0	581	1 PRLR CEREL	Q08235 cervus elap
16	216.5	8.0	625	1 TPOR_MOUSE	Q08351 mus musculus
17	215.5	7.9	581	1 PRLR_BOVIN	Q28172 bos taurus
18	215.5	7.9	625	2 Q8BRX0	Q8brx0 mus musculus
19	214.5	7.9	631	2 Q6UAN4	Q6uan4 tetraodon n
20	213	7.8	538	1 IL1R HUMAN	Q9hbes homo sapien
21	212.5	7.8	903	2 Q6ICE0	Q6ice0 homo sapien
22	211.5	7.8	635	1 TPOR HUMAN	P40238 homo sapien
23	210	7.7	529	1 IL1R_MOUSE	Q9jhx3 mus musculus
24	210	7.7	616	1 PRLR_RABIT	P14787 corytolagus
25	208.5	7.7	889	2 Q6NSJ8	Q6nsj8 homo sapien
26	208.5	7.7	897	1 CYRB_HUMAN	P32927 homo sapien
27	207.5	7.6	469	2 Q8C2G1	Q8c2g1 mus musculus
28	206	7.6	467	2 Q63216	Q63216 rattus norv
29	202.5	7.5	581	1 PRLR SHEEP	Q46561 ovis aries
30	202	7.4	468	1 IL9R_MOUSE	Q01114 mus musculus
31	201	7.4	884	2 Q6UAM6	Q6uam6 tetraodon n

32 195 7.2 622 2 Q9N0J7
33 192.5 7.1 890 2 Q9Z1A0
34 191.5 7.1 610 1 PRLR RAT
35 189 7.0 896 2 Q64146
36 189 7.0 896 2 Q782F5
37 188.5 6.9 614 2 Q62IY6
38 187 6.9 819 2 Q616F7
39 186.5 6.9 896 1 CYRB_MOUSE
40 184.5 6.8 622 1 PRLR_HUMAN
41 184 6.8 831 1 PRLR_CHICK
42 183.5 6.8 830 1 PRLR_COLLI
43 182.5 6.7 831 2 Q6QDA0
44 182 6.7 896 2 Q8QZX9
45 180 6.6 608 1 PRLR_MOUSE

ALIGNMENTS

RESULT 1
EPOR HUMAN STANDARD; PRT; 508 AA.
AC P19235; Q15443;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Erythropoietin receptor precursor (EPO-R).
GN Name=EPOR;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC TISSUE=Erythrocyte, and Fetal liver;
RX MEDLINE=90304340; PubMed=2163696;
RA Jones S.S., D'Andrea A.D., Haines L.L., Wong G.G.;
RT "Human erythropoietin receptor: cloning, expression, and biologic
characterization.";
RL Blood 76:31-35(1990).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC TISSUE=Placenta;
RX MEDLINE=92399733; PubMed=1668606;
RA Noguchi C.T., Bae K.S., Chin K., Wada Y., Schechter A.N.,
Hankins W.D.;
RT "Cloning of the human erythropoietin receptor gene.";
RL Blood 78:2548-2556(1991).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC TISSUE=Erythrocyte;
RX MEDLINE=91372359; PubMed=1654273;
RA Ehrenman K., St John T.;
RT "The erythropoietin receptor gene: cloning and identification of
multiple transcripts in an erythroid cell line OCIM1.";
RL Exp. Hematol. 19:973-977(1991).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS EPOR-F; EPOR-S AND EPOR-T).
RC TISSUE=Bone marrow, and Megakaryoblast;
RX PubMed=1324524;
RA Nakamura Y., Komatsu N., Nakauchi H.;
RT "A truncated erythropoietin receptor that fails to prevent programmed
cell death of erythroid cells.";
RL Science 257:1138-1141(1992).
RN [5]
RP SEQUENCE OF 1-96 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92399734; PubMed=1668607;
RA Maouche L., Tournamille C., Hattab C., Boffa G., Cartron J.-P.,
Chretien S.;
RT "Cloning of the gene encoding the human erythropoietin receptor.";
RL Blood 78:2557-2563(1991).
RN [6]


```

CC      Isotd=PI9235-1; Sequence=Displayed;
CC      Name=EPOR-S; Synonyms=Soluble form;

Query Match      100.0%; Score 2715; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 4.8e-181;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MDHLGASLPQVGSLLCLLAGAAPPENLPDPKFESKAALLAARGPEELLCFTERLEDL 60
Db      1 MDHLGASLPQVGSLLCLLAGAAPPENLPDPKFESKAALLAARGPEELLCFTERLEDL 60

Qy      61 VCFWEEAASAGVGNYSFSLQDEDEPKLCRLHQAPRTARGAVFWSCLPTADTSSSFVPL 120
Db      61 VCFWEEAASAGVGNYSFSLQDEDEPKLCRLHQAPRTARGAVFWSCLPTADTSSSFVPL 120

Qy      121 ELRVTASGAPRYHRVTHINEVLLDAPVGLVARLADESHGVLRWLPPPTMTSHIRY 180
Db      121 ELRVTASGAPRYHRVTHINEVLLDAPVGLVARLADESHGVLRWLPPPTMTSHIRY 180

Qy      181 EVDVSAGNAGSVORVEILEGRTECVLSNLGRTRYTFVAVRMAEPSPFGFWSAWSEPV 240
Db      181 EVDVSAGNAGSVORVEILEGRTECVLSNLGRTRYTFVAVRMAEPSPFGFWSAWSEPV 240

Qy      241 SLLTPSDLDPLILTLILVILVILVILVILVILVILVILVILVILVILVILVILVILVILV 300
Db      241 SLLTPSDLDPLILTLILVILVILVILVILVILVILVILVILVILVILVILVILVILVILV 300

Qy      301 KGNFQLWLYQNDGCLWSPCTPFTEDPPASLEVLSEKWCMTQAVEPCTDDEGLLSPVG 360
Db      301 KGNFQLWLYQNDGCLWSPCTPFTEDPPASLEVLSEKWCMTQAVEPCTDDEGLLSPVG 360

Qy      361 SEHAQDTYLVLDKWLPRNPESDLPGFGSVDIVAMDEGSEASSCSALASKPSPEGAS 420
Db      361 SEHAQDTYLVLDKWLPRNPESDLPGFGSVDIVAMDEGSEASSCSALASKPSPEGAS 420

Qy      421 AASPEYITLDPSSQLLRPWLCPRLPTPHLKYLYLVWSDSGISTDYSSGDSOGAGGL 480
Db      421 AASPEYITLDPSSQLLRPWLCPRLPTPHLKYLYLVWSDSGISTDYSSGDSOGAGGL 480

Qy      481 SDGYSNPYENSLIPAAEPLPPSVVACS 508
Db      481 SDGYSNPYENSLIPAAEPLPPSVVACS 508

RESULT 2
EPOR_PIG      STANDARD;      PRT;      509 AA.
AC      QSMY29;
DT      29-MAR-2004 (Rel. 43, Created)
DT      29-MAR-2004 (Rel. 43, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Erythropoietin receptor precursor (EPO-R).
GN      Name=EPOR;
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RX      PubMed=10962196; DOI=10.1016/S0739-7240(00)00062-X;
RA      Pearson P.L., Smith T.P.L., Sonstegard T.S., Klemcke H.G.,
RA      Christenson R.K., Vallet J.L.;
RT      "Porcine erythropoietin receptor: molecular cloning and expression in
RT      embryonic and fetal liver.";
RL      Domest. Anim. Endocrinol. 19:25-38(2000).
CC      -1- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC      induced erythroblast proliferation and differentiation. Upon EPO
CC      stimulation, EPOR dimersizes triggering the JAK2/STAT5 signaling
CC      cascade. In some cell types, can also activate STAT1 and STAT3.
CC      May also activate LYN tyrosine kinase (By similarity).
CC      -1- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC      phosphorylated form interacts with several SH2 domain-containing

```

```

CC      proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
CC      JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
CC      domain of PTPN6 binds Tyr-455 and inhibits signaling through
CC      dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
CC      signaling. Binding to PTPN11, preferentially through the N-
CC      terminal SH2 domain, promotes mitogenesis and phosphorylation of
CC      PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
CC      surface expression. Interaction with the ubiquitin ligase NODIP
CC      mediates EPO-induced cell proliferation. Interacts with ATXN2L (By
CC      similarity).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- DEVELOPMENTAL STAGE: Low expression at day 24 gestation in fetal
CC      liver. Expression increases dramatically thereafter to day 30.
CC      Levels then remain constant up to day 40.
CC      -1- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC      folding and thereby efficient intracellular transport and cell-
CC      surface receptor binding.
CC      -1- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC      activation.
CC      -1- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to
CC      as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This
CC      motif is involved in modulation of cellular responses. The
CC      phosphorylated ITIM motif can bind the SH2 domain of several SH2-
CC      containing phosphatases.
CC      -1- PTM: On EPO stimulation, phosphorylated on C-terminal tyrosine
CC      residues by JAK2. The phosphotyrosine motifs are also recruitment
CC      sites for several SH2-containing proteins and adapter proteins
CC      which mediate cell proliferation. Phosphorylation on Tyr-455 is
CC      required for PTPN6 interaction. Tyr-427 for PTPN11. Tyr-427 is
CC      also required for SOCS3 binding, but Tyr-455/Tyr-457 motif is the
CC      preferred binding site (By similarity).
CC      -1- PTM: Ubiquitinated by NODIP; appears to be either multi-
CC      monoubiquitinated or polyubiquitinated. Ubiquitination mediates
CC      proliferation and survival of EPO-dependent cells (By similarity).
CC      -1- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC      -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).
CC      -----
CC      EMBL; AF274305; AAF77065.1; -.
CC      HSSP; PI9235; IERN.
CC      InterPro; IPR002996; Cytokn_recept_B/G.
CC      InterPro; IPR009167; EPO_receptor.
CC      InterPro; IPR003961; FN_III.
CC      InterPro; IPR008957; FN_III-like.
CC      InterPro; IPR003528; Hemtrecept_L_F1.
CC      Pfam; PF00041; fn3; 1.
CC      PIRSF; PIRSF001959; EPO_receptor; 1.
CC      SMART; SM00060; FN3; 1.
CC      PROSITE; PS00853; FN3; 1.
CC      PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
CC      Glycoprotein; Phosphorylation; Receptor; Signal; Transmembrane;
CC      Ubl conjugation.
CC      SIGNAL      1      24      Potential.
CC      CHAIN      25      509      Erythropoietin receptor.
CC      DOMAIN      25      251      Extracellular (Potential).
CC      TRANSMEM      252      274      Potential.
CC      DOMAIN      275      509      Cytoplasmic (Potential).
CC      DOMAIN      145      241      Fibronectin type-III.
CC      SITE      117      117      Required for ligand binding (By
CC      similarity).
CC      SITE      234      238      WSXWS motif.
CC      SITE      282      291      Box 1 motif.
CC      SITE      369      369      APS binding, and STAT5 binding and
CC      activation (By similarity).
CC      SITE      427      427      Required for STAT5/PTPN11/SOCS3 binding.

```

```

FT SITE 453 458 ITIM motif.
FT SITE 455 455 PTPN6 binding.
FT DISULFID 52 62 By similarity.
FT DISULFID 91 107 By similarity.
FT MOD_RES 369 369 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 427 427 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 455 455 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 457 457 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 469 469 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 486 486 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 490 490 Phosphotyrosine (by JAK2) (By similarity).
FT MOD_RES 505 505 Phosphotyrosine (by JAK2) (By similarity).
FT CARBOHYD 184 184 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 509 AA; 55183 MW; 35B565D07C6BCD8A CRC64;

Query Match 84.1%; Score 2283.5; DB 1; Length 509;
Best Local Similarity 84.5%; Pred. No. 6e-151;
Matches 430; Conservative 17; Mismatches 61; Indels 1; Gaps 1;

QY 1 MDHICASLPQVGSICLLLAGAAWPPNLPDPKFSKAAALLAARGPELLCFTLRLEDL 60
DB 1 MYHFGATLPWGVSLCLLAGTAWPSPNDPAKFESKAAALLAARGPELLCFTLRLEDL 60
QY 61 VCFWEAASAGVPGNYSFSYQLEDEPMKLCRLHQARTAGAVRWCMSLPTADTSSFFVPL 120
DB 61 VCFWEAASAGVPGNYSFSYQLEDEPMKPCHLHQGTARTAGSVRWCMSLPTADTSSFFVPL 120
QY 121 ELRVT-AASGAPRYHVRHINEVLLDAPVGLVARLADSGHVLRWLWLPPTPTMTSHIR 179
DB 121 ELRVTSSGAPRYHRIHINEVLLDPPAGLLARRAESGHVLRWLWLPFPFAPNASLIR 180
QY 180 YEVDSVAGNAGSVORVILEGRTCVLSNLGRTRYTFVAVRMAWEPFSGFWSANSEP 239
DB 181 YEVNISTENAGGVORVILDRTECVLSNLGRTRYTFVWRMAWEPFSGFWSANSEP 240
QY 240 VSLTPSDLPILTLTSLILVLLVLLTALLSHRRALKQKINWGPSPSEPEGLFTT 299
DB 241 ASLTLASDLPLTLTSLILVLLVLLTALLSHRRTLKQKINWGPSPSEPEGLFTT 300
QY 300 HKNPQLWLYQNDGCLWSPCTPFTEDPPASLEVLVSERCWGTQMVAVERGTDDEGFLLEPV 359
DB 301 HKNPQLWLYQTDGCLWSPCTPFAEDPPAPLEVLVSERCWGTQAVEPAADDEGSLLEPV 360
QY 360 GSEHAQDTYLVLDKWLPRNPSEDLPGGSGVDIVAMDEGSEASSCSALASKPSPEGA 419
DB 361 GSEHARDTYLVLDKWLPRNPSEDLPGGSDLDNAAWDEASEAFCSALALKPGPEGA 420
QY 420 SAASFEYTLDPSSQLLRPWLTCPELPTTPHLKYLVLVWSDSGISTDYSSGDSQGAQG 479
DB 421 SAASFEYTLDPSSQLLRPWLTCPELPTTPHLKYLVLVWSDSGISTDYSSGDSQGTGG 480
QY 480 LSDGYSNPYENSLIPAEPLPPSVVACS 508
DB 481 SSSGYSNPYENSLVPAPEPSPPNVTCS 509

RESULT 3
EPOR_MOUSE
ID EPOR_MOUSE STANDARD; PRT; 507 AA.
AC P14753; Q63852;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Erythropoietin receptor precursor (EPO-R).
GN Name=Epor;

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_taxID=10090;
RN [1] SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RP TISSUE=Erythrocytoblasts;
RC MEDLINE=89195238; PubMed=2539263; DOI=10.1016/0092-8674(89)90965-3;
RA D'Andrea A.D., Lodish H.F., Wong G.G.;
RT "Expression cloning of the murine erythropoietin receptor.";
RL Cell 57:277-285(1989).
RN [2] SEQUENCE FROM N.A. (ISOFORMS EPOR-F AND EPOR-S).
RP STRAIN=BALE/c; TISSUE=Erythrocytoblasts, and Liver;
RC MEDLINE=91080149; PubMed=2175360;
RA Kuramochi S., Ikawa Y., Todokoro K.;
RT "Characterization of murine erythropoietin receptor genes.";
RL J. Mol. Biol. 216:567-575(1990).
RN [3] SEQUENCE FROM N.A. (ISOFORM EPOR-F), AND ALTERNATIVE SPLICING DUE TO
RP FRIEND SPLEEN FOCUS-FORMING VIRUS.
RC TISSUE=Erythrocytoblasts;
RX MEDLINE=92017832; PubMed=1656233;
RA Hino M., Tojo A., Misawa Y., Morii H., Takaku F., Shibuya M.;
RT "Unregulated expression of the erythropoietin receptor gene caused by
RT insertion of spleen focus-forming virus long terminal repeat in a
RT murine erythrocytoblast cell line.";
RL Mol. Cell. Biol. 11:5527-5533(1991).
RN [4] SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RP STRAIN=C57BL/6J, and FVB/N-3; TISSUE=Breast tumor;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5] SEQUENCE OF 1-27 FROM N.A.
RP MEDLINE=90281750; PubMed=2162479;
RA Youssoufian H., Zon L.I., Orkin S.H., D'Andrea A.D., Lodish H.F.;
RT "Structure and transcription of the mouse erythropoietin receptor
RT gene.";
RL Mol. Cell. Biol. 10:3675-3682(1990).
RN [6] SEQUENCE OF 1-24 FROM N.A.
RP MEDLINE=91201346; PubMed=1849897;
RA Lacombe C., Chretien S., Lemarchandel V., Mayeux P., Romeo P.H.,
RA Gisselbrecht S., Cartron J.-P.;
RT "Spleen focus-forming virus long terminal repeat insertion
RT activation of the murine erythropoietin receptor gene in the T3C1-2
RT friend leukemia cell line.";
RL J. Biol. Chem. 266:6952-6956(1991).
RN [7] FUNCTION IN ERYTHROPOIETIN-INDUCED MITOGENESIS, AND MUTAGENESIS OF
RP GLN-304; TRP-306; SER-317; GLU-324; LEU-330 AND GLU-331.
RP MEDLINE=93180826; PubMed=8382775;
RA Miura O., Cleveland J.L., Ihle J.N.;

```


RT "Inactivation of erythropoietin receptor function by point mutations
 in a region having homology with other cytokine receptors.";
 Mol. Cell. Biol. 13:1788-1795(1993).
 [8]
 RP FUNCTION OF THE WSXWS MOTIF, AND MUTAGENESIS OF TRP-232; SER-233;
 ALA-234; TRP-235 AND SER-236.
 PubMed=8617735; DOI=10.1074/jbc.271.9.4699;
 RA Hilton D.J., Watowich S.S., Katz L., Lodish H.F.;
 RT "Saturation mutagenesis of the WSXWS motif of the erythropoietin
 receptor.";
 J. Biol. Chem. 271:4699-4708(1996).
 [9]
 RP INTERACTION WITH JAK2, PHOSPHORYLATION, AND MUTAGENESIS OF TRP-306.
 PubMed=8068943;
 RA Miura O., Nakamura N., Quelle F.W., Witthuhn B.A., Ihle J.N., Aoki N.;
 RT "Erythropoietin induces association of the JAK2 protein tyrosine
 kinase with the erythropoietin receptor in vivo.";
 Blood 84:1501-1507(1994).
 [10]
 RP INTERACTION WITH PTPN6, AND MUTAGENESIS OF TYR-453; 453-TYR--TYR-455
 AND TYR-455.
 PubMed=7889566; DOI=10.1016/0092-8674(95)90351-8;
 RA Klingmüller U., Lorenz U., Cantley L.C., Neel B.G., Lodish H.F.;
 RT "Specific recruitment of SH-PTP1 to the erythropoietin receptor causes
 inactivation of JAK2 and termination of proliferative signals.";
 Cell 80:729-738(1995).
 [11]
 RP INTERACTION WITH PTPN11, AND MUTAGENESIS OF TYR-367; TYR-425; TYR-453;
 TYR-455; TYR-467; TYR-484; TYR-488 AND TYR-503.
 PubMed=8639815;
 RA Tauchi T., Damen J.E., Toyama K., Feng G.-S., Broxmeyer H.B.,
 Krysstal G.;
 RT "Tyrosine 425 within the activated erythropoietin receptor binds Syt,
 reduces the erythropoietin required for Syt tyrosine phosphorylation,
 and promotes mitogenesis.";
 Blood 87:4495-4501(1996).
 [12]
 RP INTERACTION WITH STAT5.
 PubMed=8665851;
 RA Gobert S., Chretien S., Gouilleux F., Muller O., Pallard C.,
 Dusanter-Fourt I., Groner B., Lacombe C., Gieselbrecht S., Mayeux P.;
 RT "Identification of tyrosine residues within the intracellular domain
 of the erythropoietin receptor crucial for STAT5 activation.";
 EMBO J. 15:2434-2441(1996).
 [13]
 RP INTERACTION WITH STAT5, AND MUTAGENESIS OF ARG-153; GLN-304; TRP-306;
 SER-317; GLU-324; LEU-330; GLU-331 AND TYR-367.
 PubMed=8657137;
 RA Quelle F.W., Wang D., Nosaka T., Thierfelder W.E., Stravopodis D.,
 Weinstein Y., Ihle J.N.;
 RT "Erythropoietin induces activation of Stat5 through association with
 specific tyrosines on the receptor that are not required for a
 mitogenic response.";
 Mol. Cell. Biol. 16:1622-1631(1996).
 [14]
 RP INTERACTION WITH CRKL AND LYN.
 MEDLINE=21413951; PubMed=11443118; DOI=10.1074/jbc.M102924200;
 RA Arai A., Kanda E., Nosaka Y., Miyasaka N., Miura O.;
 RT "Crkl is recruited through its SH2 domain to the erythropoietin
 receptor and plays a role in Lyn-mediated receptor signaling.";
 J. Biol. Chem. 276:33282-33290(2001).
 [15]
 RP INTERACTION WITH APS.
 MEDLINE=22510236; PubMed=12444928; DOI=10.1042/BJ20020716;
 RA Wollberg P., Lennartsson J., Gottfridsson E., Yoshimura A.,
 Ronnstrand L.;
 RT "The adapter protein APS associates with the multifunctional docking
 sites Tyr-568 and Tyr-936 in c-Kit.";
 Biochem. J. 370:1033-1038(2003).
 [16]
 RP PHOSPHORYLATION, AND STAT5 ACTIVATION.
 PubMed=11290583; DOI=10.1182/blood.V97.8.2230;
 RA Barber D.L., Beattie B.K., Mason J.M., Nguyen M.H.-H., Yoakim M.,

RA Neel B.G., D'Andrea A.D., Frank D.A.;
 RT "A common epitope is shared by activated signal transducer and
 activator of transcription-5 (STAT5) and the phosphorylated
 erythropoietin receptor: implications for the docking model of STAT
 activation.";
 Blood 97:2230-2237(2001).
 [17]
 RP C-GLYCOSYLATION, MUTAGENESIS OF SER-233 AND ALA-234, AND
 CARBOHYDRATE-LINKAGE SITE TRP-232.
 PubMed=12859190; DOI=10.1021/bi034112p;
 RX Furmanek A., Hess D., Rogniaux H., Hofsteenge J.;
 RA "The WSXWS motif is C-hexosylated in a soluble form of the
 erythropoietin receptor.";
 Biochemistry 42:8452-8458(2003).
 CC -!- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
 induced erythroblast proliferation and differentiation. Upon EPO
 stimulation, EPOR dimerizes triggering the JAK2/STAT5 signaling
 cascade. In some cell types, can also activate STAT1 and STAT3.
 CC May also activate the LYN tyrosine kinase.
 CC -!- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
 phosphorylated form interacts with several SH2 domain-containing
 proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
 JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
 domain of PTPN6 binds Tyr-453 and inhibits signaling through
 dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
 signaling. Binding to PTPN11, preferentially through the N-
 terminal SH2 domain, promotes mitogenesis and phosphorylation of
 PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
 surface expression. Interaction with the ubiquitin ligase NOSIP
 mediates EPO-induced cell proliferation. Interacts with ATRXN2L (by
 similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1).
 CC Secreted (isoform 2).
 CC -!- ALTERNATIVE PRODUCTS.
 CC Event-Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=EPOR-F; Synonyms=Membrane-bound form;
 CC IsoID=P14753-1; Sequence=Displayed;
 CC Name=EPOR-S; Synonyms=Soluble form;
 CC IsoID=P14753-2; Sequence=VSP_009512; VSP_009513;
 CC -!- TISSUE SPECIFICITY: Expressed in relatively mature erythroid
 progenitor cells and in EPO-responsive erythroleukemia cells.
 CC -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
 folding and thereby efficient intracellular transport and cell-
 surface receptor binding.
 CC -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
 activation.
 Query Match 80.1%; Score 2174.5; DB 1; Length 507;
 Best Local Similarity 81.7%; Pred. No. 2.4e-143;
 Matches 415; Conservative 24; Mismatches 68; Indels 1; Gaps 1;
 QY 1 MDHLGASLWPQVGSICLLLAGAANAPPNLPDPKFSKAAALLAARGPBELLCTERLEDL 60
 DB 1 MDKURVLPWPVRVGLCLLLAGAAWAPSPSLPDPKFSKAAALLASRSEELLCTORLEDL 60
 QY 61 VCFVEEASAGVPGNYSFYSQLEDEPKLCRLHOAPTARGAVRWCWSLPTADTSSFVPL 120
 DB 61 VCFVEEASAGVPGNYSFYSQLEDEPKLCRLHOAPTARGAVRWCWSLPTADTSSFVPL 119
 QY 121 ELRVTAASGAPRYHRVTHINEVLLDAPVGLVARLADESHVVLRLWLPPTPTMTSHIRY 180
 DB 120 ELQVTEASGSPRYHRIIHNIEVLLDAPAGLLARRAESGSHVLRWLWLPPTGAPMTTHIRY 179
 QY 181 EVDVSAGNGASGVORVELLEGTECVLSNLGRTRYTFAVRARMAEPFSGFWSAWSEPV 240
 DB 180 EVDVSAGNRAGGTQORVEVLEGRTECVLSNLGRGTRYTFAVRARMAEPFSGFWSAWSEPA 239
 QY 241 SLLTPSLDPLILTLTLVVLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLV 300
 DB 240 SLLTASDLPLILTLTLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLVLLV 299
 QY 301 KGNFQLWLYQNDGCLWMSPTCTFTTDDPPASLEVLSERCWGTMQAVPEGTDDGPLEPVG 360

```

Db 300 KGNFQMLLQDGLWSPGSSFFEDPAHLEVLSEPRWAVTQAGDPGADDEGLLEPVG 359
Qy 361 SEHAQDTYLVLDKLLRNPPSPDLPGGSGVDIVAMDEGSEASSCSALASKSPSPGAS 420
Db 360 SEHAQDTYLVLDKLLRTPCSENLSPGSGVDIVTWDSEATSSCSFDSLAKRPRPGTS 419
Qy 421 AASFEYTLDPSSQLLPWTLCPBLPTTPHLYLVWSDSGISTDYSGDSQGAQGL 480
Db 420 PSSPEYTLDPSSQLLPWTLCPBLPTTPHLYLVWSDSGISTDYSGDSQGVHGD 479
Qy 481 SDGYSNPNYENSLPAAEPPLPPSVACS 508
Db 480 SDGYSHPYENSLVPDSEPLHPGVACS 507

RESULT 4
EPOR_RAT STANDARD; PRT; 507 AA.
AC Q07303;
DT 01-FEB-1995 (Rel. 31, Last Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Erythropoietin receptor precursor (EPO-R).
GN Name=Epor;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM EPOR-F).
RC TISSUE=Pheochromocytoma;
RX MEDLINE=93266574; PubMed=7684373;
RA Masuda S., Nagao M., Takahata K., Konishi Y., Gallyas F., Tabira T.,
RA Sasaki R.;
RT "Functional erythropoietin receptor of the cells with neural
RT characteristics. Comparison with receptor properties of erythroid
RT cells."
RL J. Biol. Chem. 268:11208-11216(1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS EPOR-F AND EPOR-S).
RC STRAIN=Long Evans;
RX PubMed=9029168; DOI=10.1016/S0304-3835(96)04544-2;
RA Fujita M., Takahashi R., Kitada K., Watanabe R., Kitazawa S.,
RA Ashoori F., Liang P., Saya H., Serikawa T., Maeda S.;
RT "Alternative splicing of the erythropoietin receptor gene correlates
RT with erythroid differentiation in rat hematopoietic and leukemic
RT cells."
RL Cancer Lett. 112:47-55(1997).
CC -1- FUNCTION: Receptor for erythropoietin. Mediates erythropoietin-
CC induced erythroid proliferation and differentiation. Upon EPO
CC stimulation, EPOR dimerizes triggering the JAK2/STAT5 signaling
CC cascade. In some cell types, can also activate STAT1 and STAT3.
CC May also activate LYN tyrosine kinase (By similarity).
CC -1- SUBUNIT: Forms homodimers on EPO stimulation. The tyrosine-
CC phosphorylated form interacts with several SH2 domain-containing
CC proteins including LYN, the adapter protein APS, PTPN6, PTPN11,
CC JAK2, PI3 kinases, STAT5A/B, SOCS3 and CRKL. The N-terminal SH2
CC domain of PTPN6 binds Tyr-453 and inhibits signaling through
CC dephosphorylation of JAK2. APS binding also inhibits the JAK-STAT
CC signaling. Binding to PTPN11, preferentially through the N-
CC terminal SH2 domain, promotes mitogenesis and phosphorylation of
CC PTPN11. Binding of JAK2 (through its N-terminal) promotes cell-
CC surface expression. Interaction with the ubiquitin ligase NOD1P
CC mediates EPO-induced cell proliferation. Interacts with ATRXN2L (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=EPOR-F; Synonyms=Full-length form;
CC IsoId=Q07303-1; Sequence=Displayed;
CC Name=EPOR-S; Synonyms=Soluble form;
CC IsoId=Q07303-2; Sequence=VSP_009514, VSP_009515;

```

```

CC -1- TISSUE SPECIFICITY: Both isoforms expressed in bone marrow, spleen
CC and erythroleukemia cell lines.
CC -1- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC folding and thereby efficient intracellular transport and cell-
CC surface receptor binding.
CC -1- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC activation.
CC -1- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred to
CC as the immunoreceptor tyrosine-based inhibitor motif (ITIM). This
CC motif is involved in modulation of cellular responses. The
CC phosphorylated ITIM motif can bind the SH2 domain of several SH2-
CC containing phosphatases.
CC -1- PTM: On EPO stimulation, phosphorylated on C-terminal tyrosine
CC residues by JAK2. The phosphotyrosine motifs are also recruitment
CC sites for several SH2-containing proteins and adapter proteins
CC which mediate cell proliferation. Phosphorylation on Tyr-453 is
CC required for PTPN6 interaction. Tyr-425 for PTPN11. Tyr-425 is
CC also required for SOCS3 binding, but Tyr-453/Tyr-455 motif is the
CC preferred binding site (By similarity).
CC -1- PTM: Ubiquitinated by NOD1P; appears to be either multi-
CC monoubiquitinated or polyubiquitinated. Ubiquitination mediates
CC proliferation and survival of EPO-dependent cells (By similarity).
CC -1- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC Subfamily 1.
CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D13566; BAA02761.1; --
CC PIR: A46713; A46713.
CC HSSP: P19235; 1EBA.
CC RGD: 2560; Epor.
CC InterPro: IPR002996; Cytokn_recept_B/G.
CC InterPro: IPR009167; EPO_receptor.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR008957; FN_III-like.
CC InterPro: IPR003528; HemptreceptL_Fl.
CC Pfam: PF00041; fn3; 1.
CC PROSITE: PIRSFO01959; EPO_receptor; 1.
CC PROSITE: PS00853; FN3; 1.
CC PROSITE: PS01352; HEMATOPO_REC_L_F1; 1.
CC Alternative splicing: Glycoprotein; Phosphorylation; Receptor; Signal;
CC Transmembrane; Ubl conjugation.
CC SIGNAL 1 24 By similarity.
CC CHAIN 25 507 Erythropoietin receptor.
CC DOMAIN 25 249 Extracellular (Potential).
CC TRANSMEM 250 272 Potential.
CC DOMAIN 273 507 Cytoplasmic (Potential).
CC DOMAIN 143 239 Fibronectin type-III.
CC SITE 116 116 Required for ligand binding (By
CC similarity).
CC SITE 232 236 WSXWS motif.
CC SITE 281 289 Box 1 motif.
CC SITE 367 367 APS binding, and STAT5 binding and
CC activation.
CC SITE 425 425 Required for STAT5/PTPN11/SOCS3 binding.
CC SITE 451 456 ITIM motif.
CC SITE 453 453 PTPN6 binding.
CC DISULFID 52 62 By similarity.
CC DISULFID 90 106 By similarity.
CC MOD_RES 367 367 Phosphotyrosine (by JAK2) (By
CC similarity).
CC MOD_RES 425 425 Phosphotyrosine (by JAK2) (By
CC similarity).
CC MOD_RES 453 453 Phosphotyrosine (by JAK2) (By
CC similarity).
CC MOD_RES 455 455 Phosphotyrosine (by JAK2) (By
CC similarity).

```

RA	David B., Lim G.B., Moritz K.M., Koukoulas I., Wintour M.;	
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AY029232; AAK9737.1; -	
DR	HSP; P19235; IEBA.	
DR	GO; GO:0004872; F:receptor activity; IEA.	
DR	InterPro; IPR003961; FN III.	
DR	InterPro; IPR008957; FN_III-like.	
DR	InterPro; IPR003528; HemptreceptL_F1.	
DR	Pfam; PF00041; fn3; 1.	
DR	SMART; SW00060; FN3; 1.	
DR	PROSITE; PS0853; FN3; 1.	
DR	PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.	
KW	Receptor.	
FT	NON_TER 1	
FT	NON_TER 387 387	
SQ	SEQUENCE 387 AA; 42039 MW; 0D1E6173C432EBC6 CRC64;	
Query Match 61.5%; Score 1670.5; DB 2; Length 387;		
Best Local Similarity 82.9%; Pred. No. 2.4e-108;		
Matches 319; Conservative 14; Mismatches 51; Indels 1; G		
Qy	68 ASAGVGPNGNYSFSQLEDEPWKLCRLHQAPARGAVRFWCSLPTADTSSFVPLELR	
Db	3 ATAGVGPDNYSFSQLEGEKPKCRLHOTPTARGLYRFWCSLPTADTSSFVPLELR	
Qy	128 -SGAPRYHRVTHINEVLLDAPVGLVARLADSEGHVVLRWLPPPTMTSHIRYEY	
Db	63 SSGASRYRTHINEVLLDPPARLVARRADEGHVVLRWLPPPGAPMASLIRYEY	
Qy	187 GNGAGSVORVELSGRTCEVLNLRGTRTTFVARRARMAEPSEFGGFWSAWSEPSVLS	
Db	123 ENAAGGQRVILDRGTCELLNLRGGTRYTFVARRARMAEPSEFGGFWSAWSEPSALS	
Qy	247 DLDPLILTLILVILVLLVTLALLSHRRALKOKIWPGPSPESEFEGFLTTHKGKG	
Db	183 DLDPLILTLILVILVLLALLAVALLSHRRTLKOKIWPGPSPESEFEGFLTTHKG	
Qy	307 WLYQNDGCLMWSPTPTFPDPPASLEVLSRCHGTQAVEPGTDDGGLLEPVGSE	
Db	243 WLYQTGDCMWSPTFPDPPAPLEVLSECCWGVQAVEPGADDGGLLEPVGSE	
Qy	367 TYVLVDKWLPRNPSEDLPGPGSGVDIVAMDEGSEASSCSSALAKSPGECASANA	
Db	303 SYIMLDKWLPRSPSEDLQPCGDMDIVTTDSEASSCTSAALKPGPEGASANA	
Qy	427 TILDPSQLLRPMTLCPLEPTPPH 451	
Db	363 TILDPSQLLRPALPPELLPTPPH 387	
RESULT 6		
Q95N14		
ID	Q95N14 PRELIMINARY; PRT; 418 AA.	
AC	Q95N14;	
DT	01-DEC-2001 (TReMBLrel. 19, Created)	
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)	
DT	01-MAR-2004 (TReMBLrel. 26, Last annotation update)	
DE	Erythropoietin receptor (Fragment).	
OS	Ovis aries (Sheep).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;	
OC	Caprinae; Ovis.	
OX	NCBI_TaxID=9940;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	David B., Lim G.B., Moritz K.M., Koukoulas I., Wintour M.;	
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	
RL	EMBL; AY029231; AAK38170.1; -	
DR	HSP; P19235; IEBA.	
DR	GO; GO:0004872; F:receptor activity; IEA.	
DR	InterPro; IPR003961; FN III.	
DR	InterPro; IPR008957; FN_III-like.	
DR	InterPro; IPR003528; HemptreceptL_F1.	

[illegible]

Matches 194; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 59 DLVCFWEEAASAGVPGNYSFYSQLEDEPWKLCRLHQAPARGAVRWCSLPTADTSSVF 118
 Db 1 DLVCFWEEAATAGVPGDNYSFYSQLEGEFPWPCRLHQAPTARGLVRFWCSLPTADTSSVF 60

QY 119 PLELRVTAASGAPRYHRVHIHNEVLLDAPVGLVARLADSGHVLRWLPPTPTMTSHI 178
 Db 61 PLELHVTAASGASRRHTIHVNEVLLDPPARLVARRADEGHHVLRWLPPTGAPMASLI 120

QY 179 RYEVDSAGNAGSVQRYVEILLEGTECVLSNLRGRTRYTFVARRMAEPSPFGFWASWE 238
 Db 121 RYEVNISAENAAGSAQRVEILDGTECLSNLRGGTRYTFVARRMAEPSPFGFWASWE 180

QY 239 PVSLTTPSDPLDPLTLTSLIIVLVLLTVLALLSHRRALKOKIWPPIP 287
 Db 181 PASLLTASDLPLTLTSLVLLVLLALLSHRRTLKOKIWPPIP 229

RESULT 9

Q28206 PRELIMINARY; PRT; 229 AA.

AC Q28206;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Erythropoietin receptor (Fragment).
 GN Name=EpoR;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]

SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA Sullivan H.B., Feldman B. F., Majiwa P.A.O., Logan-Henfrey L.L.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U61399; AAB03871.1; -.
 DR HSSP; P19235; IEBA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR002996; Cytkn recept_B/G.
 DR InterPro; IPR003961; FN_III-like.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003528; HemtreceptL_F1.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS0853; FN3; 1.
 DR PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
 DR Receptor.
 KW NON_TER 1
 FT NON_TER 229
 SQ SEQUENCE 229 AA; 25170 MW; EDFAA6F110D992E8 CRC64;

Query Match 37.2%; Score 1011; DB 2; Length 229;
 Best Local Similarity 84.3%; Pred. No. 1.3e-62;
 Matches 193; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 59 DLVCFWEEAASAGVPGNYSFYSQLEDEPWKLCRLHQAPARGAVRWCSLPTADTSSVF 118
 Db 1 DLVCFWEEAATAGVPGDNYSFYSQLEGEFPWPCRLHQAPTARGLVRFWCSLPTADTSSVF 60

QY 119 PLELRVTAASGAPRYHRVHIHNEVLLDAPVGLVARLADSGHVLRWLPPTPTMTSHI 178
 Db 61 PLELHVTAASGASRRHTIHVNEVLLDPPARLVARRADEGHHVLRWLPPTGAPMASLI 120

QY 179 RYEVDSAGNAGSVQRYVEILLEGTECVLSNLRGRTRYTFVARRMAEPSPFGFWASWE 238
 Db 121 RYEVNISAENAAGSAQRVEILDGTECLSNLRGGTRYTFVARRMAEPSPFGFWASWE 180

QY 239 PVSLTTPSDPLDPLTLTSLIIVLVLLTVLALLSHRRALKOKIWPPIP 287
 Db 181 PASLLTASDLPLTLTSLVLLVLLALLSHRRTLKOKIWPPIP 229

RESULT 10

Q6UAP7 PRELIMINARY; PRT; 529 AA.

AC Q6UAP7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Class I helical cytokine receptor number 9.
 GN Name=CRFA9;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontoidea; Tetraodontidae; Tetraodon.
 OC NCBI_TaxID=99883;
 RN [1]

SEQUENCE FROM N.A.
 RP Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castell V., Poulain M., Vacherie B.,
 RA Blimont C., Skalli Z., Cattolico L., Poulain J., Berardinis Vd.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Wolff J.-N., Guigs R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quittier F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Crolius H.R.;
 RT "Analysis of the Tetraodon nigroviridis genome reveals the
 RT protokaryotype of bony vertebrates and its duplication in teleost
 RT fish."
 RL Nature 0:0-0(2004).
 DR EMBL; AX374481; AAR25672.1; -.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR002996; Cytkn recept_B/G.
 DR InterPro; IPR003961; FN_III-like.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003528; HemtreceptL_F1.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 DR PROSITE; PS0853; FN3; 1.
 DR PROSITE; PS01352; HEMATOPO_REC_L_F1; UNKNOWN_1.
 DR Receptor.
 KW RECEPTOR.
 SQ SEQUENCE 529 AA; 59609 MW; C4A04BFC94E9C3B6 CRC64;

Query Match 19.5%; Score 529.5; DB 2; Length 529;
 Best Local Similarity 27.8%; Pred. No. 1.4e-28;
 Matches 152; Conservative 84; Mismatches 199; Indels 111; Gaps 16;

QY 15 LCLLLAGAAWAPPNLPDPK-FESKAALLAARGPELLCFTLERLEDLVCFWEEAASAGVG 73
 Db 10 LLLTIIFGATPPANVOGAQDFRKKVIMLKEDPKPKFAEGRKDFICFWEDEERAGS 69

QY 74 PONYSFYSQLEDEPWKLCRLHQAPARGAVRWCSLPTADTSSVFPLELRVTAASGAPRY 133
 Db 70 VDOYTFYAYQNSNSRCPKLSISADSKRFLFICHLNR--IKMFQMDIQV-HREGMLIH 126

QY 134 HRVTHINEVLLDAPVGLVARLADSGHVLRWLPPTPTMTSHIRIYEYDVDSAGNAGSV 193
 Db 127 NRSLLVEMVFLDPPANNVTNTRKQGLNVTWVPPLKYMDDSMWYEVSYSAM--SHV 184

QY 194 QRVLEIGRTECVLSNLRGRTRYTFVARRMAEPSPFGFWASWEPSVLLT-PSLDPLI 252
 Db 185 MQEVMVQASSLLILRGLQPGTKYEQVRVKLDGISYSGYWSAWSVVIETLPABLDLII 244

QY 253 LTLILVILVLLTVLALLSHRRALKOKIWPPIPSPSESEFGLTTHKGNFOLWYOND 312
 Db 245 VSLALVILVILGLFLTTVMNSNRRYLVKKIWNPIPTDSKFGHGLFSVYGGFQSWLEQTS 304

QY 313 GCLMWSPTPTFTEDPPASLEVLSRRCMGTMQAVEFGTDDEGFLLEP----- 358

Db 305 G-LWTPVFFNTEELASLEVLSELCPSPSPAPLPDPEDKLPAAPNCVTD 355
 Qy 359 -----VGSEHAQDTYLVLDKWLPP-----RNP----- 380
 Db 356 GLVERRKVEANSVSEGVETADNQMPDTSWRGPQPCVPCSRPLLESQDAYVTLSNN 415
 Qy 381 -----PSEDLPGFGGSDIVAMDEGSEASSCSALASKPSPEG---ASAAASFEYT 427
 Db 416 QREBEPLNHIPEETLP-----IEKFTSRKQICESHSDLGSMQOQPASSHLSSQSFE-- 468
 Qy 428 ILDPSSQLLRWTLCPELPPPHLPKYLVLVVSIGISTDYSSGDSQGAQGLSDGYSN 487
 Db 469 -----CPNY--TWMSKGYVHMAVDSGVSMDS-----PMQVRAD-VHNN 505
 Qy 488 PYENSL 493
 Db 506 EYKNGI 511

RESULT 11

IL9R HUMAN STANDARD; PRT: 522 AA.
 AC Q01113; Q14634; Q8WUL; Q96TF0;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Interleukin-9 receptor precursor (IL-9R).
 GN Name=IL9R; Synonyms=IL9R;
 GN and
 GN Name=IL9RY; Synonyms=IL9R;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=92302307; PubMed=1376929;
 RX Renauld J.C., Druet C., Kermouni A., Houssiau F., Uyttenhove C.,
 RA van Roost E., van Snick J.;
 RT "Expression cloning of the murine and human interleukin 9 receptor
 cDNAs";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5690-5694 (1992).
 [2]
 RP MEDLINE=94250901; PubMed=8193355;
 RX Chang M.S., Engel G., Benedict C., Basu R., McNinch J.;
 RA "Isolation and characterization of the human interleukin-9 receptor
 gene";
 RL Blood 83:3199-3205 (1994).
 [3]
 RP MEDLINE=96115587; PubMed=8666384;
 RX Kermouni A., van Roost E., Arden K.C., Vermeesch J.R., Weiss S.,
 RA Godelaine D., Flint J., Lurquin C., Szikora J.P., Higgs D.R.,
 RA Marynen P., Renauld J.C.;
 RT "The IL-9 receptor gene (IL9R): genomic structure, chromosomal
 localization in the pseudoautosomal region of the long arm of the sex
 chromosomes, and identification of IL9R pseudogenes at 9qter, 10pter,
 RT 16pter, and 18pter";
 RL Genomics 29:371-382 (1995).
 [4]
 RP MEDLINE=2012249; PubMed=10655549; DOI=10.1093/hmg/9.3.395;
 RX Ciccodicola A., D'Esposito M., Esposito T., Gianfrancesco F.,
 RA Migliaccolo C., Miano M.G., Marazzito M.R., Vacca M., Franze A.,
 RA Cucurese M., Cocchia M., Curci A., Terracciano A., Torino A.,
 RA Cocchia S., Mercadante G., Pannone E., Archidiacono N., Rocchi M.,
 RA Schlesinger D., D'Urso M.;
 RT "Differentially regulated and evolved genes in the fully sequenced
 RT Xq/yq pseudoautosomal region";
 RL Hum. Mol. Genet. 9:395-401 (2000).

SEQUENCE FROM N.A.
 Rieder M.J., Amel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
 Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
 "SeattleSNPs: NHLBI HL66682 program for genomic applications, UM-
 FHCR, Seattle, WA (URL: http://pga.gs.washington.edu)";
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: This is a receptor for interleukin-9.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein and secreted.
 CC -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
 folding and thereby efficient intracellular transport and cell-
 surface receptor binding.
 CC -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
 activation.
 CC -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
 CC -!- SIMILARITY: Subfamily 4.
 CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See http://www.ebi.ac.uk/announcements/
 or send an email to license@ebi.ac.uk).

EMBL; M84747; AAA58679.1; -;
 EMBL; S71404; AAB30844.2; ALT_SEQ.
 EMBL; S71420; AAD14081.1; -;
 EMBL; L39064; AAC29513.1; -;
 EMBL; AJ271736; CAB96817.1; -;
 EMBL; AY071830; AAL55435.1; -;
 PIR; B45268; B45268.
 Genew; HGNC:6030; IL9R.
 MIM; 300007; -;
 GO; GO:0005615; C:extracellular space; TAS.
 GO; GO:0005887; C:integral to plasma membrane; TAS.
 GO; GO:0004319; F:interleukin-9 receptor activity; TAS.
 GO; GO:0008283; P:cell proliferation; TAS.
 GO; GO:0007165; P:signal transduction; TAS.
 InterPro; IPR002996; Cytokn recept_B/G.
 InterPro; IPR008957; FN_III-like.
 InterPro; IPR003531; Hemtreceptus_F1.
 PROSITE; PS01355; HEMATOPO_REC_SF1; 1.
 Glycoprotein; Polymorphism; Receptor; Signal; Transmembrane.
 SIGNAL 1 40
 CHAIN 41 522 Interleukin-9 receptor.
 DOMAIN 41 270 Extracellular (Potential).
 TRANSMEM 271 291 Potential.
 DOMAIN 292 522 Cytoplasmic (Potential).
 DOMAIN 150 244 Fibronectin type-III.
 DOMAIN 429 439 Poly-Ser.
 DOMAIN 440 443 Poly-Asn.
 SITE 245 249 WSXWS motif.
 SITE 301 309 Box 1 motif.
 CARBOHYD 117 117 N-linked (GLNAC...) (Potential).
 CARBOHYD 156 156 N-linked (GLNAC...) (Potential).
 VARIANT 239 239 E -> Q (in dbSNP:6522).
 FTID=VAR_014804.
 G -> R (in Ref. 1 and 2).
 CONFLICT 331 331 Missing (in Ref. 3 and 4).
 CONFLICT 439 439 Missing (in Ref. 3 and 4).
 SQ SEQUENCE 522 AA; 57233 MW; BBB73D6E2FAE37CB CRC64;

Query Match 9.2%; Score 251; DB 1; Length 522;
 Best Local Similarity 25.9%; Pred. No. 3.5e-09;
 Matches 112; Conservative 48; Mismatches 164; Indels 108; Gaps 22;
 Qy 45 RGPET-LICFTERLEDDLVCFWEASAGVGPNGYSFVLEDEPWKLCRLHQAP----- 97
 Db 47 QGPRSTFTCLTNILRIDCHW-SAPELGQ-----SSPMLLFTSNQAPGTHK 94
 Qy 98 -TARGAVRWCWSLPTADTSFVPLE-LRVT---AAG-----APRYHRVHINE 141

Db 95 CILRSE---CTVLPPEAVLPSDNFTTITHHCKMSGREQVSLVDPEYLPRRH----- 144
Qy 142 VLLDAPVGLVARLADESHVYLRW-LPPETPTMTSHIRYEDVDSAGNGA-GSVQRVEIL 199
Db 145 -VKLPPSDLQSNIS--SGHCLTWSISPALEPTWLLSYELAFKQEEAEQAQRDHI 201
Qy 200 EGRTECVLSNLGRTRYTFVAR-----MAEPSFGGFWMSAEPVSLTTPSLD 249
Db 202 VGVTLILAEFELDPGFHEARLRVQMATLEDDVVEERYTCQWSENQPVCFQAPORQ 261
Qy 250 FLILTL-----SLILAVILVLT---VLLSHRRALKQKIWPGISPESEFGLPTH 300
Db 262 FLIPGWGPNLWAVSLFLLTGTYLLFKLSR--VKRIFYQNVPSPMFFOPLYSVH 319
Qy 301 KGNFQLWL-YQNDGCLMWSPT-----PFTDPPASLEVLSEKWCWGTQVAEPGTDD 351
Db 320 NGNFQTMGAHGAGVLLSQDCAGTPOGALFCVQEATALLTCGPAPKWSVALEE---EQ 376
Qy 352 EGPILPEVGSQAQD-----TYVLVDKMLPRNPSEDLPGPGGSDIVAM 397
Db 377 EGPCTRLPGNLSSEDLVPAGCTEWRVQTLAYLPOEDWA---PTSLTRPAPDS----- 426
Qy 398 DEGSEASSCSA 409
Db 427 -EGSRSSSSSS 437

RESULT 12
Q61YE8
ID Q61YE8 PRELIMINARY; PRT; 630 AA.
AC Q61YE8, 2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Thrombopoietin receptor.
GN Name=mpl;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartunek P., Karafiat V., Bartunkova J., Dvorakova M., Kralova J.,
RA Pajer P., Zinke M., Dvorak M.,
RA Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY613435; AAT4555.1; -
DR GO: 0004872; P:receptor activity; IEA.
DR InterPro: IPR003961; FN_III.
DR Pfam: PF00041; fn3; 1.
DR SMART: SM00060; FN3; 2.
DR PROSITE: PS00853; FN3; 1.
KW Receptor.
SQ SEQUENCE 630 AA; 68856 MW; 9E8FCC6CE61E0EDE CRC64;

Query Match 8.5%; Score 231.5; DB 2; Length 630;
Best Local Similarity 26.6%; Pred. No. 1e-07;
Matches 71; Conservative 32; Mismatches 107; Indels 57; Gaps 9;

Qy 24 WAP-----PPNLPDPKFKESKAALLAARGPEELLCFTELEDVLCFWEESAASAG 71
Db 13 WLPALLTAVLLSRHSPTADTEPVTSDAALL-AEVPEDILCFSRSPEDLTCFWEDEASA 71
Qy 72 VGGNYSFSYQLEDEPKLCRLHQAPTARGAVFWCSLPTADTSSVPLELRV-TAASGA 130
Db 72 I-----HRFIYW--DAP--TTCVSTQSGVSGMHCIVFSPQDVRLEFTPLHLVLDTTNR 124
Qy 131 PRYHRVTHINEVLLDAPVGLVARLADESHVYLRWLP-----PETPMT 175
Db 125 TRHRELSDVAGLIAPPVNIARWAGAAQLRVSWQPPADYPNPFLEYEVCHPMTPTV 184
Qy 176 -----SHIRYEDVDSAGNGAGSVQRVEILEGRTECVLSNLGRTRYTF 218

Db 185 TPRGTAPGDLPARPTGRAHPPTARGAASQGTGO-----GLVQADTWVVLRELQPCVRVHI 240
Qy 219 AVARMAEPSPGFGFWMSAEPVSLTTP 245
Db 241 QVRSKPDGTSMDGVGWPVSEVVVAETP 267

RESULT 13
Q925F5
ID Q925F5 PRELIMINARY; PRT; 566 AA.
AC Q925F5, 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE C-mpl-II.
GN Name=Mpl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21992449; PubMed=11997107; DOI=10.1016/S0167-4781(01)00357-8;
RA Sabath D.F., Lofton-Day C., Lin N., Lok S., Kenneth Kaushansky.,
RA Broudy V.C.;
RA "Identification and characterization of an isoform of murine Mpl.";
RL Biochim. Biophys. Acta 1574:383-386(2002).
DR EMBL: AF360122; AAK52492.1; -
DR HSSP: PI9235; IERN.
DR MGD: MGI:97076; Mpl.
DR GO: 0005615; C:extracellular space; TAS.
DR GO: 0016021; C:integral to membrane; TAS.
DR Pfam: PF00041; fn3; 1.
DR SMART: SM00060; FN3; 1.
DR PROSITE: PS00853; FN3; 1.
DR PROSITE: PS01352; HEMATOPO RECL F1; UNKNOWN 1.
DR PROSITE: PS00041; HTH ARAC FAMILY 1; UNKNOWN 1.
SQ SEQUENCE 566 AA; 63339 MW; A19D3DB5F7BDBA6B CRC64;

Query Match 8.3%; Score 224; DB 2; Length 566;
Best Local Similarity 21.8%; Pred. No. 3e-07;
Matches 128; Conservative 42; Mismatches 167; Indels 250; Gaps 23;

Qy 16 CULLAGAAWAPPNLPD-PKFESKAALLAARGPEELLCFTELEDVLCFWEESAASAGVGP 74
Db 5 CULLA-----LPNQAQVTSQDVFLALGTEPLNCFSTQFEDLTCFWEDEEAA--PS 53
Qy 75 GNYFSYQLEDEPKLCRLHQAPTARGAVFWCSLPTAD-TSFPVPLELRVTAAS-GAPR 132
Db 54 GTYQLLYAYRGEKPRACPLYSQSVPTFGTRYVCQPPAQDEVRLEFFPLHLWKNVSLNQT 113
Qy 133 YHRVTHI-----NEVLLDAP----- 148
Db 114 IQRVLFVDSVETCCPTLMMNPVPVLDQPPCVHTASQPHGPVRTSPAGEAPFLTVKGS 173
Qy 149 ---VGLVA-----RLADESHGVLR--WLP----- 168
Db 174 CLVSGLOAGKSWLQRLSQDQGVLSRGSGWSPFPVTVLDLPGDAVTIGLCQFTLDLKMVT 233
Qy 169 -----PPETPM-----TSH 177
Db 234 CQWQQDRTSSQGFHRSRTRCCPTDRDPTWEKCEEEPRPGSQPALVSRCHFKSRNSV 293
Qy 178 IRYEDVDSAGNGA-----GSVQRVEILEGRTE----- 204
Db 294 IHLVEVTTAQAGVHSLGSPFIHQAVLLPTFSLHWRVSGRLELEWHQOSSWAAQET 353
Qy 205 C-----VLS-----NLGRTRYTFVARMAEPSPGFGFWMSAEPV 240
Db 354 CYQLRYTGEGRDVKVLEPSLGARGGTLRLPRARYSLQLRARLNGPTTQGPNSAWSPPA 413
Qy 241 SLLTSPDLDPLILTLISLILVILVLLTVLALL-----SHRRALKQKIWPGISPESEFEG 295


```

Db 414 RVSTGSETAWITVALLVLSALLGLLLKQFPAHYRRLRHAWPSLPD----- 466
Qy 296 LFTTHKGNFQWLYQNDCLWWSPTPTED-----PPASLEVLSERCWGT---MQAVEP 347
Db 467 ---LHR---VLGQYLRDTAA--LSPSKATVTDSCVEEPSLLEILPKSSESPLPLCPSPQ 519
Qy 348 GTDDEG-----PLLEPVGS-----EHAQDTVLVDKMLLP 377
Db 520 QMDYRGLQCLRTMPLSVCPMAETGSCCTTHANHSVLPISYMQP 566

RESULT 14
PRLR_PIG
ID PRLR_PIG STANDARD; PRT; 625 AA.
AC Q6JTA8; O46679; Q6JTA9; Q6JTB0;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Prolactin receptor precursor (PRL-R).
GN Name=PRLR;
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RC STRAIN=Landrace, and Yucatan;
RA Trott J.F., Hovey R.C.;
RT "Cloning and characterization of multiple forms of the pig (Sus
RT scrofa) prolactin receptor.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 531-625 FROM N.A.
RX MEDLINE=97468290; PubMed=9321486;
RA Vincent A.L., Wang L., Tuggle C.K., Robic A., Rothschild M.F.;
RL "Prolactin receptor maps to pig chromosome 16.";
RM Mann, Genome 8:793-794(1997).
CC -!- FUNCTION: This is a receptor for the anterior pituitary hormone
CC prolactin (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The WSXWS motif appears to be necessary for proper protein
CC folding and thereby efficient intracellular transport and cell-
CC surface receptor binding (By similarity).
CC -!- DOMAIN: The box 1 motif is required for JAK interaction and/or
CC activation.
CC -!- SIMILARITY: Belongs to the type I cytokine family of receptors.
CC subfamily 1.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY308824; AAQ76841.1; -
CC EMBL; AY308825; AAQ76842.1; -
CC EMBL; AY308826; AAQ76843.1; -
CC EMBL; U96306; AAC02897.1; -
CC InterPro; IPR002996; Cytokn_recept_B/G.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR003528; HemptreceptL_F1.
CC Pfam; PF00041; fn3; 2.
CC SMART; SM00060; FN3; 2.
CC PROSITE; PS00853; FN3; 2.
CC PROSITE; PS01352; HEMATOPO_REC_L_F1; 1.
KW Glycoprotein; Receptor; Repeat; Signal; Transmembrane.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 625 Prolactin receptor.

```

```

FT DOMAIN 25 237 Extracellular (Potential).
FT TRANSMEM 238 258 Potential.
FT DOMAIN 259 625 Cytoplasmic (Potential).
FT DOMAIN 27 123 Fibronectin type-III 1.
FT DOMAIN 127 227 Fibronectin type-III 2.
FT DOMAIN 111 114 Poly-Ser.
FT SITE 215 219 WSXWS motif.
FT SITE 267 275 Box 1 motif.
FT DISULFID 36 46 By similarity.
FT CARBOHYD 75 86 By similarity.
FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 104 104 N-linked (GlcNAc...) (Potential).
FT CONFLICT 132 132 P -> R (in Ref. 1; AAQ76842).
FT CONFLICT 372 372 L -> P (in Ref. 1; AAQ76842).
FT CONFLICT 406 406 A -> G (in Ref. 1; AAQ76842).
FT CONFLICT 461 461 K -> R (in Ref. 1; AAQ76842).
FT CONFLICT 480 480 M -> L (in Ref. 1; AAQ76842).
FT CONFLICT 510 510 G -> S (in Ref. 1; AAQ76842).
FT CONFLICT 534 534 S -> G (in Ref. 1; AAQ76842).
FT CONFLICT 597 597 S -> G (in Ref. 1; AAQ76842).
SQ SEQUENCE 625 AA; 69572 MW; 751B2B593ABFC51B CRC64;

Query Match 8.1%; Score 220; DB 1; Length 625;
Best Local Similarity 23.8%; Pred. No. 6.4e-07; Indels 117; Gaps 25;
Matches 113; Conservative 65; Mismatches 179;

Qy 27 PPNLPDPKPKESKAALLAARGPEELLCTFERLEDLVCFWEAAASAGVGPNGYSFYSQLEDE 86
Db 27 PPKRPE-----IFKRSPEK-----EFTCMKPKGADGGL-PTNYLTLYHKEGE 69
Qy 87 PWKLCRLHQAPTARGAVRWCSLPTADTSFVPLELRVTA-----ASGAPRHYRHHIN 140
Db 70 TF---THECPDYKTGPNSCYFNKKHTSIWTIYITVATNQMGSSSDPRY---VDVT 122
Qy 141 EVVLLDAPVGLVARL---ADESCHVVLRLPPTETPMTS---HRYEVDVSNAGGASVQ 194
Db 123 YIVEPDPVNNLTLELKKPKDQKPLKWLPPTLVDVRSGLWTLQVEIRLOPEKTA--EW 180
Qy 195 RVEILGRTECVLSNLRGRTRYFAVRARMAEBSFGFMSAWSEPVSLTLPDPLILT 254
Db 181 ETHFAGQOTFKILSLYPGQKYLQVRCK--PDH-GFWSEWSPSSIQIPNDFSMKDDT 236
Qy 255 LSIL-----VTVLLTVALLSHRRALKQKIWPQIPSPSEFEGFLFTH---KGNFQL 306
Db 237 MWIFVAILSAAVCLIMIAVALKGY--SMVACILPPVPGP--KIKG-FDTHLLEKKGK-- 288
Qy 307 WLQNDGCLWWSPTPTPTEDPPASLEVLSE-RCWGTMQAVEPGTDDEGPLEPVGSEHAQ 365
Db 289 -----SEILLSALGC-----QDFPPTSDEDLLEVELEVDDSE 321
Qy 366 DTVLVDKMLLRNPPSEDLPGFGSVDIVAMDEGSEASCS--ALASKP- 414
Db 322 -----DQQLMPAH--SKEHPSQGRKPTHLDPDSDSGRGSCDSPSLSEKCEPRANPPK 373
Qy 415 --SPEGASAAAFYITLDSSQLLRWTLCPELPPTPHLKLYLVVSDSGIST 466
Db 374 FHTPEGIEKFG-----DPETNLRP-----QDPQSTSVESKLLLYFHADGSKSST 417

RESULT 15
PRLR_CEREL
ID PRLR_CEREL STANDARD; PRT; 581 AA.
AC Q28235;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Prolactin receptor precursor (PRL-R).
GN Name=PRLR;
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervinae; Cervus.
OC NCBI_TaxID=9860;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 04:08:54 ; Search time 189 Seconds
(without alignments)
1039.546 Million cell updates/sec

Title: US-09-016-159D-5

Perfect score: 2715
Sequence: 1 MDHIGASLWPQVGLCLLLA.....YENSLPAEPLPPSYVACS 508

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2715	100.0	508	2 AAR06512	Aar06512 EPO recep
2	2715	100.0	508	2 AAR70032	Aar70032 Human ery
3	2715	100.0	508	2 AAR69503	Aar69503 Human ery
4	2715	100.0	508	8 ADO05724	Ado05724 Human ery
5	2715	100.0	508	8 ADP12417	Adp12417 Protein e
6	2715	100.0	508	8 ABM82037	Abm82037 Tumour-as
7	2708	99.7	508	5 ABB09173	Abb09173 Human ery
8	2693	99.2	508	8 ADM93422	Adm93422 Human NOV
9	2691	99.1	508	7 ADE28677	Ade28677 Human NOV
10	2690	99.1	508	2 AAR47518	Aar47518 Human EPO
11	2407	88.7	458	7 ADE28675	Ade28675 Human NOV
12	2407	88.7	458	8 ADM93420	Adm93420 Human NOV
13	2404	88.5	458	7 ADE28671	Ade28671 Human NOV
14	2404	88.5	458	7 ADE28673	Ade28673 Human NOV
15	2404	88.5	458	8 ADM93416	Adm93416 Human NOV
16	2404	88.5	458	8 ADM93418	Adm93418 Human NOV
17	2330	85.8	438	3 AAY44622	Aay44622 Truncated
18	2322	85.5	438	3 AAY44623	Aay44623 RI54C tru
19	2171.5	80.0	507	2 AAR06511	Aar06511 EPO recep
20	2169.5	79.9	507	2 AAR47517	Aar47517 MEL EPO r
21	2169.5	79.9	507	2 AAR69502	Aar69502 Mouse ery
22	2167.5	79.8	507	2 AAR50327	Aar50327 Mouse sol
23	1352.5	49.8	503	3 AAB13012	Aab13012 Q-tagged
24	1335	49.2	488	2 AAW08349	Aaw08349 EporPC fu
25	1194	44.0	225	3 AAB21685	Aab21685 Human mat

26	1112	41.0	211	3 AAB21686	Aab21686 Human mat
27	1063.5	39.2	265	2 AAR50326	Aar50326 Mouse sol
28	371.5	13.7	107	8 ADL92633	Adl92633 IMABIS019
29	370.5	13.6	107	8 ADL92637	Adl92637 IMABIS028
30	370.5	13.6	107	8 ADR41983	Adr41983 VAP relat
31	370.5	13.6	107	8 ADR41979	Adr41979 VAP relat
32	366	13.5	108	8 ADL92635	Adl92635 IMABIS025
33	366	13.5	108	8 ADR41981	Adr41981 VAP relat
34	354.5	13.1	117	3 AAY94338	Aay94338 Human cel
35	332.5	12.2	109	8 ADL92634	Adl92634 IMABIS020
36	332.5	12.2	109	8 ADR41980	Adr41980 VAP relat
37	325.5	12.0	109	8 ADL92636	Adl92636 IMABIS027
38	325.5	12.0	109	8 ADR41982	Adr41982 VAP relat
39	286.5	10.6	109	8 ADL92535	Adl92535 IMAB1100
40	286.5	10.6	109	8 ADR41849	Adr41849 Verbatille
41	275.5	10.1	109	8 ADL92527	Adl92527 IMab600 V
42	275.5	10.1	109	8 ADR41842	Adr41842 Verbatille
43	261.5	9.6	500	2 AAW64057	Aaw64057 Human IL-
44	251	9.2	501	2 AAW64055	Aaw64055 Human IL-
45	251	9.2	521	6 AAO19700	Aao19700 IL-13 mod

ALIGNMENTS

RESULT 1
AAR06512
ID AAR06512 standard; protein; 508 AA.

XX AC AAR06512;

XX XX 25-MAR-2003 (revised)

DT 04-JAN-1991 (first entry)

XX XX EPO receptor.

XX XX Erythropoietin; Diamond Blackfan anaemia; polycythemia vera.

XX OS Homo sapiens.

XX PN WO9008822-A.

XX PD 09-AUG-1990.

XX PF 03-FEB-1989; 89US-00306503.

XX PR 03-FEB-1989; 89US-00306503.

XX (GEM) GENETICS INST INC.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Dandrea A, Wong G;

XX WPI; 1990-260931/34.

XX N-PSDB; AAQ05748.

XX Erythropoietin receptor and gene - used for developing reagents and systems to control and study erythropoiesis.

XX Disclosure; Fig 2; 53pp; English.

XX The sequence was deduced from DNA obt'd. from a clone isolated from a commercially available human genomic cDNA library in phage Lambda Fix (Stratagene). The sequence encodes a type I trans- membrane protein with binding affinity for EPO. The gene and recombinant EPO receptor produced on expression of the DNA are used to develop reagents and systems to control and study erythropoiesis. It is believed that the EPO receptor is dys- functional in individuals with Diamond Blackfan anaemia, and may be hyperactive in polycythemia vera. See also AAR06511 (murine EPO receptor). (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 508 AA;

SQ

```

Query Match      100.0%; Score 2715; DB 2; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDHLGASLWPGVGSCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
DB 1 MDHLGASLWPGVGSCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

QY 61 VCFWEAASAGVGNYSFSYQLEDEPKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120
DB 61 VCFWEAASAGVGNYSFSYQLEDEPKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120

QY 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPTPMTSHIRY 180
DB 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPTPMTSHIRY 180

QY 181 EVDVSAGNAGSVORVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFGGFWSAWSEPV 240
DB 181 EVDVSAGNAGSVORVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFGGFWSAWSEPV 240

QY 241 SLTTPSDLDPLILTLTLVLLVILVLLVTLVLLSHRRALKQKIWPGISPESEFGLFTTH 300
DB 241 SLTTPSDLDPLILTLTLVLLVILVLLVTLVLLSHRRALKQKIWPGISPESEFGLFTTH 300

QY 301 KGNFQWLWYQNDGCLWWSPTCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDGGLLEPVG 360
DB 301 KGNFQWLWYQNDGCLWWSPTCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDGGLLEPVG 360

QY 361 SEHAQDTYLVLDKWLPRNPPEDELPFGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420
DB 361 SEHAQDTYLVLDKWLPRNPPEDELPFGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420

QY 421 AASFEYITLDPSSQLLRPWTLCPGLPPTPHLKYLYLVVSDSGISTDYSSGDSOGAOGGL 480
DB 421 AASFEYITLDPSSQLLRPWTLCPGLPPTPHLKYLYLVVSDSGISTDYSSGDSOGAOGGL 480

QY 481 SDGYSNPYENSLIPAAEPLPPSYVACS 508
DB 481 SDGYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 2.
AAR70032
ID AAR70032 standard; protein; 508 AA.
XX AC AAR70032;
XX DT 25-MAR-2003 (revised)
XX DT 07-OCT-1995 (first entry)
XX DE Human erythropoietin receptor.
XX KW Erythropoietin receptor; extracellular domain.
XX OS Homo sapiens.
XX FT Key Location/Qualifiers
XX FT Domain 9..83
XX FT Domain /note= "extracellular domain"
XX FT Domain 25..250
XX FT Domain /note= "extracellular domain"
XX FT Misc-difference 25..29
XX FT /note= "forward primer AAQ82991 specific site"
XX FT Misc-difference 222..226
XX FT /note= "reverse primer AAQ82992 specific site"
XX PN WO9505469-A1.
XX PD 23-FEB-1995.
XX PF 15-AUG-1994; 94WO-US0009298.
XX

```

```

PR 16-AUG-1993; 93US-00106815.
PA (LEEJ/) LEE J Y.
XX Lee JY;
PI WPI; 1995-098767/13.
DR N-PSDB; AAQ82990.
XX
PT New pure human erythropoietin receptor fragment - obtd. by expression as
PT a fusion protein having a thrombin proteolytic cleavage site.
XX
PS Disclosure; Page 27-29; 42pp; English.
XX
CC The full-length erythropoietin receptor (EPO-R) is given. Extracellular
CC domains are expressed from vector plasmid pGEX-2T as fusion proteins with
CC glutathione-S-transferase. The domains are used for investigating the
CC structure of the EPO-R and for identifying factors involved in regulating
CC differentiation and proliferation mechanisms in erythroid progenitor
CC cells. They can also be used for identifying and quantitating EPO and EPO
CC -R as well as in understanding haematopoietic malignancy and some
CC cardiovascular system disorders. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 508 AA;

Query Match      100.0%; Score 2715; DB 2; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDHLGASLWPGVGSCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60
DB 1 MDHLGASLWPGVGSCLLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCFTERLEDL 60

QY 61 VCFWEAASAGVGNYSFSYQLEDEPKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120
DB 61 VCFWEAASAGVGNYSFSYQLEDEPKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120

QY 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPTPMTSHIRY 180
DB 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPTPMTSHIRY 180

QY 181 EVDVSAGNAGSVORVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFGGFWSAWSEPV 240
DB 181 EVDVSAGNAGSVORVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFGGFWSAWSEPV 240

QY 241 SLTTPSDLDPLILTLTLVLLVILVLLVTLVLLSHRRALKQKIWPGISPESEFGLFTTH 300
DB 241 SLTTPSDLDPLILTLTLVLLVILVLLVTLVLLSHRRALKQKIWPGISPESEFGLFTTH 300

QY 301 KGNFQWLWYQNDGCLWWSPTCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDGGLLEPVG 360
DB 301 KGNFQWLWYQNDGCLWWSPTCTPFTEDPPASLEVLSERCWGTMQAVEPGTDDGGLLEPVG 360

QY 361 SEHAQDTYLVLDKWLPRNPPEDELPFGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420
DB 361 SEHAQDTYLVLDKWLPRNPPEDELPFGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420

QY 421 AASFEYITLDPSSQLLRPWTLCPGLPPTPHLKYLYLVVSDSGISTDYSSGDSOGAOGGL 480
DB 421 AASFEYITLDPSSQLLRPWTLCPGLPPTPHLKYLYLVVSDSGISTDYSSGDSOGAOGGL 480

QY 481 SDGYSNPYENSLIPAAEPLPPSYVACS 508
DB 481 SDGYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 3
AAR69503
ID AAR69503 standard; protein; 508 AA.
XX AC AAR69503;
XX

```

DT 25-MAR-2003 (revised)
DT 11-AUG-1995 (first entry)
DE Human erythropoietin receptor.
XX Erythropoietin receptor; anemia therapy; diagnostic.
KW Homo sapiens.
OS
FH Key Location/Qualifiers
FT Peptide 1..24 /note= "signal peptide"
FT Protein 25..508 /note= "mature protein"
FT Modified-site 76..79 /note= "N-glycosylation site"
FT Domain 251..272 /note= "transmembrane region"
FT
XX
PN US5378808-A.
XX
XX 03-JAN-1995.
PD
XX
PF 10-JUN-1993; 93US-00075069.
XX
XX 03-FEB-1989; 89US-00306503.
PR 25-MAR-1991; 91US-00678877.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jones SS, D'andrea A, Wong GG;
PI WPI; 1995-051310/07.
XX N-PSDB; AAQ81892.
DR
DR recombinant erythropoietin receptor polypeptide(s) - used for detection,
PT purificn. and therapy and for prodn. of antibodies for detection and
PT therapy.
XX
XX Claim 2; Fig 9; 24pp; English.
PS
XX The sequence is that of a 55-kDa human erythropoietin receptor. The
CC receptor polypeptide may be used in purification and detection of
CC erythropoietin, and in production of antibodies for anemia therapy. The
CC polypeptide may also be used for treating individuals who are
CC hypersensitive to erythropoietin or who have elevated erythropoietin
CC levels. They may be used in therapy of e.g. primary or secondary
CC proliferative polycythemia. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 508 AA;

Query Match 100.0%; Score 2715; DB 2; Length 508;
Best Local Similarity 100.0%; Pred. No. 1.9e-216;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDHLGASLWPGVSGSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELICFTERLEDL 60
Db 1 MDHLGASLWPGVSGSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELICFTERLEDL 60
Qy 61 VCFWEEAASAGVGPNGYSFVSQLEDEPWKLCRLHQAPTARGAVRWCSLPTADTSSFVPL 120
Db 61 VCFWEEAASAGVGPNGYSFVSQLEDEPWKLCRLHQAPTARGAVRWCSLPTADTSSFVPL 120
Qy 121 ELRVTVAAGAPRYRHVHINEVLLDAPVGLVARLADESGHVWLRLWLPPTPTMTSHIRY 180
Db 121 ELRVTVAAGAPRYRHVHINEVLLDAPVGLVARLADESGHVWLRLWLPPTPTMTSHIRY 180
Qy 181 EVDVSAGNAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFSGFGFWASWSEPV 240
Db 181 EVDVSAGNAGSVQRVEILEGRTECVLSNLRGRTRYTFAVRARMAEPFSGFGFWASWSEPV 240
Qy 241 SLLTPSDLDPLILTLVLVLLVLLTVLALLSHRRALKQKIPGIPSPSESEFGLFTTH 300
Db 241 SLLTPSDLDPLILTLVLVLLVLLTVLALLSHRRALKQKIPGIPSPSESEFGLFTTH 300

Db 241 SLLTPSDLDPLILTLVLVLLVLLTVLALLSHRRALKQKIPGIPSPSESEFGLFTTH 300
Qy 301 KGNFQLMLYQNDGCLWMSPTPTFTEDPPASLEVLSERCWGTMQAVEPGTDDGEPVVG 360
Db 301 KGNFQLMLYQNDGCLWMSPTPTFTEDPPASLEVLSERCWGTMQAVEPGTDDGEPVVG 360
Qy 361 SEHAQDTYLVLDKWLPRNPSSDLPGPGGSDIVAMDEGSEASCSALASKPSPEGAS 420
Db 361 SEHAQDTYLVLDKWLPRNPSSDLPGPGGSDIVAMDEGSEASCSALASKPSPEGAS 420
Qy 421 AASFEYITLDPSSQLLRPWTLCPELPTPHLKYLYLVVSDSGISTDYSSGDSQAQOGL 480
Db 421 AASFEYITLDPSSQLLRPWTLCPELPTPHLKYLYLVVSDSGISTDYSSGDSQAQOGL 480
Qy 481 SDGFSNPYENSILIPAAEPLPPSYVACS 508
Db 481 SDGFSNPYENSILIPAAEPLPPSYVACS 508

RESULT 4
ADO05724
ID ADO05724 standard; protein; 508 AA.
XX
AC ADO05724;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human erythropoietin receptor (EPOR).
XX
KW T cell; antiallergic; immunosuppressive; virucide; antibacterial;
KW antiparasitic; cytostatic; gene therapy; human; receptor;
KW erythropoietin receptor.
XX
OS Homo sapiens.
XX
PN WO2004032867-A2.
XX
PD 22-APR-2004.
XX
PF 09-OCT-2003; 2003WO-US032065.
XX
PR 09-OCT-2002; 2002US-0417102P.
PR 09-OCT-2002; 2002US-0417103P.
PR 09-OCT-2002; 2002US-0417243P.
PR 18-OCT-2002; 2002US-0419575P.
PR 08-NOV-2002; 2002US-0424777P.
PR 08-NOV-2002; 2002US-0424881P.
XX
PA (TOLE-) TOLERRX INC.
XX
XX Rao P, Szymanska G;
XX WPI; 2004-340801/31.
DR N-PSDB; ADO05723.
DR GENBANK; 4557561.
XX
PT Treating a condition that benefits from modulating regulatory or effector
PT T cell function comprises administering an agent that modulates the
PT expression or activity of a gene or polypeptide (e.g. PTGER2, TGFbeta1,
PT Jagged-1 or GPR-32).
XX
PS Example; SEQ ID NO 24; 161pp; English.
XX
CC The invention relates to treating a subject having a condition that
CC benefits from modulating the balance of regulatory T cell function
CC relative to effector T cell function, or vice versa, in a subject. The
CC method involves administering an agent that modulates the expression or
CC activity of a molecule selected from PTGER2 and TGFbeta1, or Jagged-1,
CC GPR-32, CD83, CD84, CD89, serotonin R, BY55, serotonin R2C, GPR63,
CC histamine R-H4, GPR58, EPO-R, PSG-1, PSG-3, PSG-6, PSG-9, PDE-4d, and P1-
CC 3-related kinase to the subject, such that treatment occurs. The methods
CC are useful for diagnosing, preventing or treating conditions
CC characterized by a too-vigorous or weak effector T cell or regulatory T

CC cell response to antigens associated with the condition, such as in an
 CC allergic response, an autoimmune disorder, a viral infection, a microbial
 CC infection, a parasitic infection or a tumour. The present sequence
 CC represents a human erythropoietin receptor (EPOR), preferentially
 CC expressed in regulatory T cells.

XX SQ Sequence 508 AA;

Query Match 100.0%; Score 2715; DB 8; Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.9e-216;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDHLGASLWPOVGSCLLLAGAAWAPPNLPDPKESKAALLAARGPEELLCTERLEDL 60
 DB 1 MDHLGASLWPOVGSCLLLAGAAWAPPNLPDPKESKAALLAARGPEELLCTERLEDL 60

QY 61 VCFWEAASAGVPGNYSFYSYQLEDEPKLCRLHQAPTARGAVFMCSLPTADTSSFPVL 120
 DB 61 VCFWEAASAGVPGNYSFYSYQLEDEPKLCRLHQAPTARGAVFMCSLPTADTSSFPVL 120

QY 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180
 DB 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180

QY 181 EVDVSAGNGAGSVQVREILEGRTECVLSNLRGRTRYTFVAVRMAEPSPFGFWSAMSEPV 240
 DB 181 EVDVSAGNGAGSVQVREILEGRTECVLSNLRGRTRYTFVAVRMAEPSPFGFWSAMSEPV 240

QY 241 SLTTPSDLDPLILTLSLILVILVLLTVALLSHRRALKOKIWPICPSESEFGLFTTH 300
 DB 241 SLTTPSDLDPLILTLSLILVILVLLTVALLSHRRALKOKIWPICPSESEFGLFTTH 300

QY 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSERCWGTMOAVEPGTDDGGLLEPVG 360
 DB 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSERCWGTMOAVEPGTDDGGLLEPVG 360

QY 361 SEHAQDTYLVLDKWLPRNPSEDLPGPGGSDIVAMDEGSEASSCSALASKPSPEGAS 420
 DB 361 SEHAQDTYLVLDKWLPRNPSEDLPGPGGSDIVAMDEGSEASSCSALASKPSPEGAS 420

QY 421 AASFEYITLDPSSQLLRPWTLCPELPPTPHLKYLVLVSDSGISTDYSSGDSQGAQGL 480
 DB 421 AASFEYITLDPSSQLLRPWTLCPELPPTPHLKYLVLVSDSGISTDYSSGDSQGAQGL 480

QY 481 SDGYPYNSPYENSLIPAAEPLPPSYVACS 508
 DB 481 SDGYPYNSPYENSLIPAAEPLPPSYVACS 508

RESULT 5

ADP12417
 ID ADP12417 standard; protein; 508 AA.

XX AC ADP12417;

XX DT 12-AUG-2004 (first entry)

DE DE Protein encoded by mRNA of the invention #27.

XX KW transplant rejection; immune system; rheumatoid arthritis; lupus;
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

XX OS Homo sapiens.

XX PN WO2004042346-A2.

XX PD 21-MAY-2004.

XX PF 24-APR-2003; 2003WO-US012946.

XX PR 24-APR-2002; 2002US-00131831.

XX PR 20-DEC-2002; 2002US-00325899.

XX

PA (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
 PI Rosenberg S;

XX DR WPI; 2004-400724/37.

XX PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
 PT rejection, in an individual, comprises detecting the expression level of
 PT the genes.

XX PS Claim 65; SEQ ID NO 2426; 1762pp; English.

XX CC The present invention relates to diagnosing or monitoring transplant
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
 CC comprises detecting the expression level of one or more genes. The
 CC methods, system and kits are useful in diagnosing or monitoring
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
 CC islet, lung, bone marrow or stem cell transplant rejection,
 CC xenotransplant rejection or mechanical organ replacement rejection, in an
 CC individual. The method is also useful in assessing the immune status of
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
 CC viral, bacterial or fungal infection. The present sequence represents a
 CC protein that is encoded by the mRNA of the invention.

XX SQ Sequence 508 AA;

Query Match 100.0%; Score 2715; DB 8; Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.9e-216;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDHLGASLWPOVGSCLLLAGAAWAPPNLPDPKESKAALLAARGPEELLCTERLEDL 60
 DB 1 MDHLGASLWPOVGSCLLLAGAAWAPPNLPDPKESKAALLAARGPEELLCTERLEDL 60

QY 61 VCFWEAASAGVPGNYSFYSYQLEDEPKLCRLHQAPTARGAVFMCSLPTADTSSFPVL 120
 DB 61 VCFWEAASAGVPGNYSFYSYQLEDEPKLCRLHQAPTARGAVFMCSLPTADTSSFPVL 120

QY 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180
 DB 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180

QY 181 EVDVSAGNGAGSVQVREILEGRTECVLSNLRGRTRYTFVAVRMAEPSPFGFWSAMSEPV 240
 DB 181 EVDVSAGNGAGSVQVREILEGRTECVLSNLRGRTRYTFVAVRMAEPSPFGFWSAMSEPV 240

QY 241 SLTTPSDLDPLILTLSLILVILVLLTVALLSHRRALKOKIWPICPSESEFGLFTTH 300
 DB 241 SLTTPSDLDPLILTLSLILVILVLLTVALLSHRRALKOKIWPICPSESEFGLFTTH 300

QY 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSERCWGTMOAVEPGTDDGGLLEPVG 360
 DB 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSERCWGTMOAVEPGTDDGGLLEPVG 360

QY 361 SEHAQDTYLVLDKWLPRNPSEDLPGPGGSDIVAMDEGSEASSCSALASKPSPEGAS 420
 DB 361 SEHAQDTYLVLDKWLPRNPSEDLPGPGGSDIVAMDEGSEASSCSALASKPSPEGAS 420

QY 421 AASFEYITLDPSSQLLRPWTLCPELPPTPHLKYLVLVSDSGISTDYSSGDSQGAQGL 480
 DB 421 AASFEYITLDPSSQLLRPWTLCPELPPTPHLKYLVLVSDSGISTDYSSGDSQGAQGL 480

QY 481 SDGYPYNSPYENSLIPAAEPLPPSYVACS 508
 DB 481 SDGYPYNSPYENSLIPAAEPLPPSYVACS 508

RESULT 6
 ABM82037

CC polypeptide segment having an amino terminus and a carboxyl terminus, the
 CC segment having a thrombin proteolytic cleavage site at the carboxyl
 CC terminus; and (b) a second polypeptide segment consisting essentially of
 CC the purified human EPO receptor, where the second polypeptide segment is
 CC covalently coupled to the carboxyl terminus of the first polypeptide
 CC segment. The human EPO receptor polypeptide is useful for studying ligand
 CC binding to the EPO receptor and for quantitating the amounts of the EPO
 CC receptor. The polypeptide is also useful for analysing and understanding
 CC receptor structure and signal transduction. The present sequence
 CC represents human EPO receptor from the present invention
 XX
 SQ Sequence 508 AA;

Query Match 99.7%; Score 2708; DB 5; Length 508;
 Best Local Similarity 99.8%; Pred. No. 7,1e-216;
 Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDHLGASLWPGVSGCLLAGAAWAPPNLPDPKESKAALLAARGPEELLCFTLERL 60
 Db 1 MDHLGASLWPGVSGCLLAGAAWAPPNLPDPKESKAALLAARGPEELLCFTLERL 60

QY 61 VCFWEAASAGVPGNYSFVSQLEDEPWKLCRLHQAPTARGAVFWCSLPTADTSSFVPL 120
 Db 61 VCFWEAASAGVPGNYSFVSQLEDEPWKLCRLHQAPTARGAVFWCSLPTADTSSFVPL 120

QY 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180
 Db 121 ELRVTAAAGAPRYHRVHINEVLLDAPVGLVARLADESGHVLRWLPPPTMTSHIRY 180

QY 181 EVDVSAGNAGSVQRVEILEGRTECVLSNLRGRTRYTFVLRMAEPSEFGFWANSEPV 240
 Db 181 EVDVSAGNAGSVQRVEILEGRTECVLSNLRGRTRYTFVLRMAEPSEFGFWANSEPV 240

QY 241 SLLTPSDLDPLILTLILVILVLLTLVLLSHRRALKQKIWPGIPSESEFGLFTTH 300
 Db 241 SLLTPSDLDPLILTLILVILVLLTLVLLSHRRALKQKIWPGIPSESEFGLFTTH 300

QY 301 KGNFOLWLYQDGLWSPCTPFTEDPPASLEVLISERCGWTMQAVEPGTDDGFLLEPVG 360
 Db 301 KGNFOLWLYQDGLWSPCTPFTEDPPASLEVLISERCGWTMQAVEPGTDDGFLLEPVG 360

QY 361 SEHAQDTVLVDKWLPRNPSEDLPGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420
 Db 361 SEHAQDTVLVDKWLPRNPSEDLPGGSGVDIVAMDEGSEASSCSALAKSPPEGAS 420

QY 421 AASFEYTLIDPSSQLLRPWTLCPELPPPHLKYLYLVVSDSGISTDYSSGDSQGAQGG 480
 Db 421 AASFEYTLIDPSSQLLRPWTLCPELPPPHLKYLYLVVSDSGISTDYSSGDSQGAQGG 480

QY 481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508
 Db 481 SDGPYSNPYENSLIPAAEPLPPSYVACS 508

RESULT 8
 ADM93422
 ID ADM93422 standard; protein; 508 AA.
 XX
 AC ADM93422;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Human NOVX polypeptide #27.
 XX
 KW Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
 KW hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
 KW haemophilia; graft-versus-host disease; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW multiple sclerosis; diabetes; obesity; bronchial asthma;
 KW acquired immunodeficiency syndrome; AIDS; Crohn's disease;
 KW infectious disease; anorexia; immune disorder.
 XX
 OS Homo sapiens.

XX US2004067882-A1.
 FN 08-APR-2004.
 PD 05-NOV-2002; 2002US-00287971.
 XX 22-OCT-2001; 2001US-00035568.
 PR 05-NOV-2001; 2001US-0338626P.
 PR 06-NOV-2001; 2001US-0333072P.
 PR 09-NOV-2001; 2001US-0345398P.
 PR 09-NOV-2001; 2001US-0348283P.
 PR 15-NOV-2001; 2001US-0335610P.
 PR 21-NOV-2001; 2001US-03322152P.
 PR 28-NOV-2001; 2001US-0333912P.
 PR 29-NOV-2001; 2001US-00997425.
 PR 29-NOV-2001; 2001US-0334300P.
 PR 04-DEC-2001; 2001US-0336576P.
 PR 05-FEB-2002; 2002US-0354807P.
 PR 15-MAY-2002; 2002US-0380968P.
 PR 16-MAY-2002; 2002US-0381043P.
 PR 02-JUL-2002; 2002US-0393148P.
 PR 02-JUL-2002; 2002US-0393262P.
 PR 06-AUG-2002; 2002US-0401479P.
 PR 06-AUG-2002; 2002US-0401626P.
 PR 07-AUG-2002; 2002US-0401593P.
 PR 07-AUG-2002; 2002US-0401695P.
 PR 26-AUG-2002; 2002US-0406181P.
 XX
 PA (ALSO/) ALSOBROOK J P.
 PA (ALVA/) ALVAREZ E.
 PA (ANDE/) ANDERSON D W.
 PA (BARO/) BARON M.
 PA (BOLD/) BOLDOG F L.
 PA (BURG/) BURGESS C E.
 PA (CASM/) CASMAN S J.
 PA (CHAP/) CHAFOVAL A.
 PA (DHAN/) DHANABAL M.
 PA (EDIN/) EDINGER S R.
 PA (EISE/) EISEN A.
 PA (ELLE/) ELLERMAN K.
 PA (ETTE/) ETTENBERG S.
 PA (GANG/) GANGOLLI E A.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (GROS/) GROSSE W M.
 PA (GUOX/) GUO X.
 PA (HACK/) HACKETT C.
 PA (JIWW/) JI W.
 PA (KEKU/) KEKUDA R.
 PA (KHRA/) KHRAMTSOV N V.
 PA (LEPL/) LEPLEY D M.
 PA (LILL/) LI L.
 PA (MACD/) MACDOUGALL J R.
 PA (MALY/) MALYANKAR U M.
 PA (MAZU/) MAZUR A.
 PA (MCQU/) MCQUEENEY K.
 PA (MEZE/) MEZES P S.
 PA (MILL/) MILLER C E.
 PA (MILL/) MILLET I.
 PA (MISH/) MISHRA V.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENNA C E A.
 PA (PEYM/) PEYMAN J A.
 PA (RAST/) RASTELLI L.
 PA (RIEG/) RIEGER D K.
 PA (ROTH/) ROTHENBERG M E.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SMIT/) SMITHSON G.
 PA (SPAD/) SPADERNA S K.
 PA (STAR/) STARLING G.

PA (SPVT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (THER/) THERNEV V T.
PA (TWO/) TWOMLOW N.
PA (VERN/) VERNET C A M.
PA (ZERR/) ZERRHUSEN B D.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.
XX
PI Alsbrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L;
PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;
PI Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M;
PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME;
PI Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G;
PI Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM;
PI Zerrhuseen BD, Voss EZ, Zhong M;
XX
DR WPI; 2004-355303/33.
DR N-PSDB; ADM93421.
XX
XX Novel isolated NOVX polypeptide useful treating or preventing disorders
PT or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
XX Claim 2; SEQ ID NO 54; 330pp; English.
XX
CC The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The NOVX polypeptides and polynucleotides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of the sequences in a mammalian subject, and for
CC treating or preventing a pathology associated with NOVX. The
CC polypeptides, polynucleotides and antibodies that bind immunospecifically
CC to the polypeptides are useful for treating or preventing disorders or
CC syndromes such as congenital heart defects, cardiomyopathy,
CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC anorexia and immune disorders. This sequence represents a human NOVX
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 508 AA;

Query Match 99.2%; Score 2693; DB 8; Length 508;
Best Local Similarity 99.2%; Pred. No. 1.3e-214;
Matches 504; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 MDHLGASLWQVGSGLCLLLAGAAWAPPNLPDPKFSKALLAARGPEELLCFTEILEDL 60
DB 1 MDHLGASLWQVGSGLCLLLAGAAWAPPNLPDPKFSKALLAARGPEELLCFTEILEDL 60
QY 61 VCFWEAASAGVGNYSFYSQLEDEPWKLCRLHOAPTARGAVRWCSTPTADTSSFPVL 120
DB 61 VCFWEAASAGVGNYSFYSQLEDEPWKLCRLHOAPTARGAVRWCSTPTADTSSFPVL 120
QY 121 ELRVTASGAPRYHRVTHINEVLLDAPVGLVARLADESGHVLRWLPPTPTMTSHIRY 180
DB 121 ELRVTASGAPRYHRVTHINEVLLDAPVGLVARLADESGHVLRWLPPTPTMTSHIRY 180
QY 181 EVDVSAGNAGSVQVREILGRTECVLSNLRGRTRYTFAVRARMAPSPGCFWNSAEPV 240
DB 181 AVDVSAGNAGSVQVREILGRTECVLSNLRGRTRYTFAVRARMAPSPGCFWNSAEPV 240
QY 241 SLITPSDLDPLIITLSLIVVILVLTVALLSHRRALKOKIWPGPSSEFEGLFTTH 300
DB 241 SLITPSDLDPLIITLSLIVVILVLTVALLSHRRALKOKIWPGPSSEFEGLFTTH 300
QY 301 KGNFQLMWLYQNDGCLWNSPCTPTFTDPPASLEVLSERCWGTMQAVPGTDEGPLEPVG 360

Db 301 KGNFQLMWLYQNDGCLWNSPCTPTFTDPPASLEVLSERCWGTMQAVPGTDEGPLEPVG 360
QY 361 SEHAQDTYLVLDKWLPRNPSPSEDLPGPGSGVDIVAMDEGSEASSCSALASKPSPEGAS 420
Db 361 SEHAQDTYLVLDKWLPRNPSPSEDLPGPGSGVDIVAMDEGSEASSCSALASKPSPEGAS 420
QY 421 AASFEYTLDPSSQLLRPWTLCPELPPTPHLKLYLVVSDSGISTDYSSGDSOGAOGGL 480
Db 421 AASFEYTLDPSSQLLRPWTLCPELPPTPHLKLYLVVSDSGISTDYSSGDSOGAOGGL 480
QY 481 SDGYSNPYNSLIPAAEPLPPSYVACS 508
Db 481 SDGYSNPYNSLIPAAEPLPPSYVACS 508
RESULT 9
ADE28677
ID ADE28677 standard; protein; 508 AA.
XX AC ADE28677;
XX
DT 29-JAN-2004 (first entry)
XX DE Human NOV15d protein - SEQ ID 54.
XX
KW NOVX: antidiabetic; anorectic; cardiatic; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antinflammatory; dermatological;
KW antisthmatic; antilipemic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; NOV.
OS Homo sapiens.
XX
PN WO2003040330-A2.
XX
PD 15-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-US035536.
XX
PR 05-NOV-2001; 2001US-0338626P.
PR 05-DEC-2001; 2001US-0336600P.
PR 07-DEC-2001; 2001US-0338285P.
PR 12-DEC-2001; 2001US-0341346P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 20-DEC-2001; 2001US-0342592P.
PR 27-DEC-2001; 2001US-0344297P.
PR 31-DEC-2001; 2001US-0344903P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 07-AUG-2002; 2002US-0401788P.
PR 26-AUG-2002; 2002US-0406353P.
PR 31-OCT-2002; 2002US-00287971.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Alsbrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L;
PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;

Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;
Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
Vernet CAM, Zerhusen BD, Zhong M;
WPI: 2003-441555/41.
DR N-PSDB; ADE28676.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
XX Claim 1; SEQ ID NO 54; 447pp; English.
XX
XX The invention relates to a novel isolated NOVX polypeptide. The
CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
CC cardiatic, hypotensive, antiarteriosclerotic, virucide, antibacterial,
CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
CC dermatological, antiasthmatic and antilipidemic activities. The
CC polypeptides, nucleic acid molecules and antibodies may be useful for
CC treating or diagnosing diseases including metabolic disorders such as
CC diabetes and obesity, infectious diseases, anorexia, cancer,
CC cardiovascular diseases including hypertension and atherosclerosis,
CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
CC disorders, inflammatory skin disorders, asthma and dyslipidaemia
CC Furthermore, the nucleic acids and polypeptides may also be used to
CC identify molecules that modulate or inhibit neurogenesis, cell
CC differentiation and proliferation, haemopoiesis, wound healing and
CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC be used as hybridisation probes, in chromosome mapping, tissue typing,
CC preventive medicine and pharmacogenomics. The current sequence is that of
CC the human NOV protein of the invention.
XX
XX SQ Sequence 508 AA;

Query Match 99.1%; Score 2691; DB 7; Length 508;
Best Local Similarity 99.0%; Pred. No. 1.8e-214;
Matches 503; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCTERLEDL 60
Db 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCTERLEDL 60
QY 61 VCFWEAAAGVGPNGNYSFYQLEDEPWKLCRLHQAPTARGAVRWCSTPTADTSFVPL 120
Db 61 VCFWEAAAGVGPNGNYSFYQLEDEPWKLCRLHQAPTARGAVRWCSTPTADTSFVPL 120
QY 121 ELRVTAAAGAPRYHRVHINEVVLDPVGLVARLADESGHVLRWLPPTPTMTSHIRY 180
Db 121 ELRVTAAAGAPRYHRVHINEVVLDPVGLVARLADESGHVLRWLPPTPTMTSHIRY 180
QY 181 EVDVSGNAGSVQVRIIEGTECVLSNLRGRTRYTFAVRARMAEPSPFGGFWASWSPV 240
Db 181 EVDVSGNAGSVQVRIIEGTECVLSNLRGRTRYTFAVRARMAEPSPFGGFWASWSPV 240
QY 241 SLTTSDDLPLTLTSLILWLVLTLLVLTLLVLTLLVLTLLVLTLLVLTLLVLTLLVLT 300
Db 241 SLTTSDDLPLTLTSLILWLVLTLLVLTLLVLTLLVLTLLVLTLLVLTLLVLTLLVLT 300
QY 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSEKWCMTQAVPEGTDDGGLLEPVG 360
Db 301 KGNFOLWLYQNDGCLWSPCTPFTEDPPASLEVLSEKWCMTQAVPEGTDDGGLLEPVG 360
QY 361 SEHAQDTVLVDKWLPRNPSEDLPFGGSGVDIVAMDEGSEASSCSALASKSPESGAS 420
Db 361 SEHAQDTVLVDKWLPRNPSEDLPFGGSGVDIVAMDEGSEASSCSALASKSPESGAS 420
QY 421 AASFEYTLIDPSSQLLRPWLTPCPPLTPPHKYLVLVVSQSGISTDYSSGDSQGAQGL 480
Db 421 AASFEYTLIDPSSQLLRPWLTPCPPLTPPHKYLVLVVSQSGISTDYSSGDSQGAQGL 480

QY 481 SDGYSNPNYNSLIPAAEPLPPSYVACS 508
Db 481 SDGYSNPNYNSLIPAAEPLPPSYVACS 508

RESULT 10

AAR47518
ID AAR47518 standard; protein; 508 AA.
XX
AC AAR47518;
XX
DT 25-MAR-2003 (revised)
DT 24-JUN-1994 (first entry)
XX
XX Human EPO receptor.
XX Erythropoietin receptor; recombinant; murine; anaemia.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..24 /note="signal peptide"
XX Protein 25..508 /note="mature EPO receptor"
XX Region 251..272 /note="putative transmembrane domain"
XX
XX US5278065-A.
XX
XX 11-JAN-1994.
XX
XX 25-MAR-1991; 91US-00678877.
XX
XX 03-FEB-1989; 89US-00306503.
XX
XX (GENY) GENETICS INST INC.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
XX D'andrea A, Wong GG, Jones SS;
XX
XX WPI: 1994-025409/03.
XX N-PSDB; AAQ53995.
XX
XX Recombinant DNA encoding erythropoietin receptor - used to develop prods.
XX for study, treatment or diagnosis of disorders in which receptor is
XX dysfunctional.
XX
XX Disclosure; Fig 9; 24pp; English.
XX
XX Mouse erythroleukaemia (MEL) cells were used to construct a cDNA library.
XX The cDNA was used to transfect COS-1 cells and these were screened for
XX radioiodinated erythropoietin (EPO) binding to isolate cDNA encoding the
XX EPO receptor. This cDNA was used as a probe to screen a human genomic
XX cDNA library to obtain DNA encoding the human EPO receptor. The cDNA may
XX be used to study, treat or diagnose disorders in which the EPO receptor
XX is dysfunctional. The EPO receptor may also be used to raise antibodies or
XX for treating hypersensitivity to EPO or who have elevated levels of EPO.
XX The prod. is pref. used for treating anaemias, primary proliferative
XX polycythemia and secondary polycythemia. See also AAR47517. (Updated on
XX 25-MAR-2003 to correct PF field.)
XX
XX SQ Sequence 508 AA;

Query Match 99.1%; Score 2690; DB 2; Length 508;
Best Local Similarity 98.8%; Pred. No. 2.2e-214;
Matches 502; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCTERLEDL 60
Db 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFESKAALLAARGPEELLCTERLEDL 60

QY	61	VCWEEAASAGVGPGNYSFYSQLEDEPKLCRLHQAPTARGAVRWCSTLPTADTSSFVPL	120
Db	61	VCWEEAASAGVGPGNYSFYSQLEDEPKLCRLHQAPTARGAVRWCSTLPTADTSSFVPL	120
QY	121	ELRVTAASGAPRYHRVTHINEVLLDAPVGLVARLADESHVVLRLPPPTMTSHIRY	180
Db	121	ELRVTAASGAPRYHRVTHINEVLLDAPVGLVARLADESHVVLRLPPPTMTSHIRY	180
QY	181	EVDSAGNAGSAGVQVREILGRTECVLSNLRGRTRYTFVAVRMAEPSFGGFWMSSEPV	240
Db	181	EVDSAGNAGSAGVQVREILGRTECVLSNLRGRTRYTFVAVRMAEPSFGGFWMSSEPV	240
QY	241	SLTTPSDLPDLIILTLILVILVLLTVLALLSHRRALKQKIWPGIPSPESFEGLFTTH	300
Db	241	SLTTPSDLPDLIILTLILVILVLLTVLALLSHRRALKQKIWPGIPSPESFEGLFTTH	300
QY	301	KGNFQWLWYNDGCLWWSCTPTEDPPASLEVLSEKRCWGTQVAPGDDGGPLLEPVG	360
Db	301	KGNFQWLWYNDGCLWWSCTPTEDPPASLEVLSEKRCWGTQVAPGDDGGPLLEPVG	360
QY	361	SEHAQTYLVLDKWLPRNPPEDELPGGSGVDIVAMDEGSASSCSALAKSPSEGA	420
Db	361	SEHAQTYLVLDKWLPRNPPEDELPGGSGVDIVAMDEGSASSCSALAKSPSEGA	420
QY	421	AASFETILDPSSQLLRPWTLCPELPTPHLKLYLVVSDSGISTDYSSGDSQAQGL	480
Db	421	AASFETILDPSSQLLRPWTLCPELPTPHLKLYLVVSDSGISTDYSSGDSQAQGL	480
QY	481	SDGPNPNPNSLIPAAEPLPSYVACS	508
Db	481	SDGPNPNPNSLIPAAEPLPSYVACS	508

RESULT 11

ADE28675
ID ADE28675 standard; protein; 458 AA.

XX	AC	ADE28675;
XX	DT	29-JAN-2004 (first entry)
XX	DE	Human NOV15c protein - SEQ ID 52.
XX	KW	NOVX; antidiabetic; anorectic; cardiac; hypertensive;
XX	KW	antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
XX	KW	nootropic; neuroprotective; antiparkinsonian; anticonvulsant;
XX	KW	osteopathic; antiarthritic; antiinflammatory; dermatological;
XX	KW	antiasthmatic; antilipaeamic; metabolic; diabetes; obesity; infectious;
XX	KW	anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
XX	KW	neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
XX	KW	osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
XX	KW	neurogenesis; cell differentiation; proliferation; haemopoiesis;
XX	KW	wound healing; angiogenesis; Gene therapy; chromosome mapping;
XX	KW	tissue typing; human; NOV.
XX	OS	Homo sapiens.
XX	XX	WO2003040330-A2.
XX	PN	15-MAY-2003.
XX	PD	05-NOV-2002; 2002WO-US035536.
XX	PF	05-NOV-2001; 2001US-0338626P.
XX	PR	05-DEC-2001; 2001US-0336600P.
XX	PR	07-DEC-2001; 2001US-0338285P.
XX	PR	12-DEC-2001; 2001US-0341346P.
XX	PR	17-DEC-2001; 2001US-0341477P.
XX	PR	17-DEC-2001; 2001US-0341540P.
XX	PR	20-DEC-2001; 2001US-0342592P.
XX	PR	27-DEC-2001; 2001US-0344297P.
XX	PR	31-DEC-2001; 2001US-0344903P.

PR	17-APR-2002;	2002US-0373288P.
PR	15-MAY-2002;	2002US-0380981P.
PR	17-MAY-2002;	2002US-0381495P.
PR	28-MAY-2002;	2002US-0383534P.
PR	28-MAY-2002;	2002US-0383744P.
PR	29-MAY-2002;	2002US-0383829P.
PR	29-MAY-2002;	2002US-0384024P.
PR	07-AUG-2002;	2002US-0401788P.
PR	26-AUG-2002;	2002US-0406353P.
PR	31-OCT-2002;	2002US-00287971.
XX	(CURA-) CURAGEN CORP.	
XX	Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A, Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Grosse WM, Guo X, Hackett C, Ji W, Kerkuda R, Khrantsov NV, Lepley DM, Li L, Macdougall JR, Malyankar A, Mazur A, McQueeney K, Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M, Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA, Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM, Zerhusen BD, Zhong M;	
XX	WPI; 2003-441555/41.	
DR	N-PSDB; ADE28674.	
XX	New isolated NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.	
XX	Claim 1; SEQ ID NO 52; 447pp; English.	
XX	The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates, antidiabetic, anorectic, cardiac, hypotensive, antiarteriosclerotic, virucide, antibacterial, fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian, anticonvulsant, osteopathic, antiarthritic, antiinflammatory, dermatological, antiasthmatic and antilipaeamic activities. The polypeptides, nucleic acid molecules and antibodies may be useful for treating or diagnosing diseases including metabolic disorders such as diabetes and obesity, infectious diseases, anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and dyslipidaemia. Furthermore, the nucleic acids and polypeptides may also be used to identify molecules that modulate or inhibit neurogenesis, cell differentiation and proliferation, haemopoiesis, wound healing and angiogenesis, as well as in gene therapy. Finally, the nucleic acids may be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The current sequence is that of the human NOV protein of the invention.	
XX	Sequence 458 AA;	
XX	Query Match	88.7%; Score 2407; DB 7; Length 458;
XX	Best Local Similarity	89.2%; Pred. No. 6.1e-191;
XX	Matches 453; Conservative 1; Mismatches 4; Indels 50; Gaps 1;	
QY	1	MDHLGASLWPGVSGSLCLLLAGAAWAPPNLPDPKFESKAAALAAARGPEELCFTLRDL 60
Db	1	MDHLGASLWPGVSGSLCLLLAGAAWAPPNLPDPKFESKAAALAAARGPEELCFTLRDL 60
QY	61	VCFWEEAASAGVGPGNYSFYSQLEDEPKLCRLHQAPTARGAVRWCSTLPTADTSSFVPL 120
Db	61	VCFWEEAASAGVGPGNYSFYSQLEDEPKLCRLHQAPTARGAVRWCSTLPTADTSSFVPL 120
QY	121	ELRVTAASGAPRYHRVTHINEVLLDAPVGLVARLADESHVVLRLPPPTMTSHIRY 180
Db	121	ELRVTAASGAPRYHRVTHINEVLLDAPVGLVARLADESHVVLRLPPPTMTSHIRY 180
QY	181	EVDSAGNAGSAGVQVREILGRTECVLSNLRGRTRYTFVAVRMAEPSFGGFWMSSEPV 240

Db 181 EVDSAGNGAGSVQRVEILEGRTECVLSNLRGRTRYTFVTRMAEPSPFGFWSAWSEP 240
Qy 241 SLITPSDLPLITLTLILVILVLLVLLALLSHRRALKOKIWPFIPISESEFGLFTTH 300
Db 241 SLITPSDLPLITLTLILVILVLLVLLALLSHRRALKOKIWPFIPISESEFGLFTTH 300
Qy 301 KGNFQLWLYQNDGLMWSPTPTTDPASLEVLSEKWCMTQMAVPGTDDGFLLEPVG 360
Db 301 KGNFQLWLYQNDGLMWSPTPTTDPASLEVLSEKWCMTQMAVPGTDDGFLLEPVG 360
Qy 361 SEHAQDTYLVLDKWLPRNPSPEDLPCPGSGSDIVAMDEGSEASSCSALLASRSPEGAS 420
Db 361 SEHAQDTYLVLDKWLPRNPSPEDLPG----- 387
Qy 421 AASFEYITLDPSSQLLRPMTLCPPLPPTPHLKLYLVVSDSGISTDYSSGDSQGAQGG 480
Db 388 -----PWALCPPLPPTPHLKLYLVVSDSGISTDYSSGDSQGAQGG 430
Qy 481 SDGPNPNYENSLIPAAEPLPPSYVACS 508
Db 431 SDGPNPNYENSLIPAAEPLPPSYVACS 458

RESULT 12
ADM93420
ID ADM93420 standard; protein; 458 AA.
AC ADM93420;
XX
XX
DT 01-JUL-2004 (first entry)
XX
DE Human NOVX polypeptide #26.
XX
KW Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
KW hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
KW haemophilia; graft-versus-host disease; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW multiple sclerosis; diabetes; obesity; bronchial asthma;
KW acquired immunodeficiency syndrome; AIDS; Crohn's disease;
KW infectious disease; anorexia; immune disorder.
XX
OS Homo sapiens.
XX
XX US2004067882-A1.
XX
XX 08-APR-2004.
XX
XX 05-NOV-2002; 2002US-00287971.
XX
XX 22-OCT-2001; 2001US-00035568.
PR 05-NOV-2001; 2001US-0338626P.
PR 06-NOV-2001; 2001US-0333072P.
PR 09-NOV-2001; 2001US-0345398P.
PR 09-NOV-2001; 2001US-0348283P.
PR 15-NOV-2001; 2001US-0335610P.
PR 21-NOV-2001; 2001US-0332152P.
PR 28-NOV-2001; 2001US-0333912P.
PR 29-NOV-2001; 2001US-00997425.
PR 29-NOV-2001; 2001US-0334100P.
PR 04-DEC-2001; 2001US-0336576P.
PR 05-FEB-2002; 2002US-0354807P.
PR 15-MAY-2002; 2002US-0380968P.
PR 16-MAY-2002; 2002US-0381043P.
PR 02-JUL-2002; 2002US-0393148P.
PR 02-JUL-2002; 2002US-0393262P.
PR 06-AUG-2002; 2002US-0401479P.
PR 06-AUG-2002; 2002US-0401626P.
PR 07-AUG-2002; 2002US-0401593P.
PR 07-AUG-2002; 2002US-0401695P.
PR 26-AUG-2002; 2002US-0406181P.
XX
XX (ALSO/) ALSOBROOK J P.

PA (ALVA/) ALVAREZ E.
PA (ANDE/) ANDERSON D W.
PA (BARO/) BARON M.
PA (BOLD/) BOLDOG F L.
PA (BURG/) BURGESS C E.
PA (CASH/) CASHMAN S J.
PA (CHAP/) CHAPOVAL A.
PA (DHAN/) DHANABAL M.
PA (EDIN/) EDINGER S R.
PA (EISE/) EISEN A.
PA (ELLE/) ELLERMAN K.
PA (ETTE/) ETTENBERG S.
PA (GANG/) GANGOLLI E A.
PA (GERL/) GERLACH V.
PA (GORM/) GORMAN L.
PA (GROS/) GROSSE W M.
PA (GUOX/) GUO X.
PA (HACK/) HACKETT C.
PA (JIWU/) JI W.
PA (KEKU/) KEKUDA R.
PA (KHRA/) KHRAMTSOV N V.
PA (LEPL/) LEPLY D M.
PA (LILL/) LI L.
PA (MACD/) MACDOUGALL J R.
PA (MALY/) MALYANKAR U M.
PA (MAZU/) MAZUR A.
PA (MCQU/) MCQUEENY K.
PA (MEZE/) MEZES P S.
PA (MILL/) MILLER C E.
PA (MILL/) MILLET I.
PA (MISH/) MISHRA V.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENA C E A.
PA (PEYM/) PEYMAN J A.
PA (RAST/) RASTELLI L.
PA (RIEG/) RIEGER D K.
PA (ROTH/) ROTHENBERG M E.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (SMIT/) SMITHSON G.
PA (SPAD/) SPADERNA S K.
PA (STAR/) STARLING G.
PA (SPYT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (TCHE/) TCHERNEV V T.
PA (TWOI/) TWOMLOW N.
PA (VERN/) VERNET C A M.
PA (ZERR/) ZERHUSEN B D.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.
XX
XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL,
PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A,
PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L,
PI Lepley DM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
PI Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M;
PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME;
PI Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G;
PI Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM;
PI Zerhusen BD, Voss EZ, Zhong M;
XX
XX WPI; 2004-355303/33.
DR N-PSDB; ADM93419.
XX
PT Novel isolated NOVX polypeptide useful treating or preventing disorders
PT or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
PS Claim 2; SEQ ID NO 52; 330pp; English.
XX
XX The invention relates to human NOVX polypeptides and the polynucleotides

CC encoding them. The NOVX polypeptides and polynucleotides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of the sequences in a mammalian subject, and for
 CC treating or preventing a pathology associated with NOVX. The
 CC polypeptides, polynucleotides and antibodies that bind immunospecifically
 CC to the polypeptides are useful for treating or preventing disorders or
 CC syndromes such as congenital heart defects, cardiomyopathy,
 CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
 CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
 CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
 CC immunodeficiency syndrome (AIDS). Crohn's disease, infectious disease,
 CC anorexia and immune disorders. This sequence represents a human NOVX
 CC polypeptide of the invention. Note: The sequence data for this patent is
 CC also available from USPTO at seqdata.uspto.gov/sequence.html.
 XX

SQ Sequence 458 AA;

Query Match 88.7%; Score 2407; DB 8; Length 458;
 Best Local Similarity 89.2%; Pred. No. 6.1e-191;
 Matches 453; Conservative 1; Mismatches 4; Indels 50; Gaps 1;

Qy	1	MDHLGASLWPGVSGSLCLLLAGAAWAPPNLPDPKFSKAAALLAARGPEELLCTFERLEDL	60
Db	1	MDHLGASLWPGVSGSLCLLLAGAAWAPPNLPDPKFSKAAALLAARGPEELLCTFERLEDL	60
Qy	61	VCFWEEAASAGVGNYSFYSQLEDEPWKLCRLHQAPTARGVFWCSLPTADTSSFPVL	120
Db	61	VCFWEEAASAGVGNYSFYSQLEDEPWKLCRLHQAPTARGVFWCSLPTADTSSFPVL	120
Qy	121	ELRVTASGAPRHRVTHINEVLLDAPVGLVARLADESGHVLRWLPPTETMTSHIRY	180
Db	121	ELRVTASGAPRHRVTHINEVLLDAPVGLVARLADESGHVLRWLPPTETMTSHIRY	180
Qy	181	EVDSAGNAGSVQVREILGRTECVLSNLRGRTRYFAVRMAEPSPFGFSANSEPV	240
Db	181	EVDSAGNAGSVQVREILGRTECVLSNLRGRTRYFAVRMAEPSPFGFSANSEPV	240
Qy	241	SLTTPSLDLPLTLTSLVILVLLTVLALLSHRRALKQKIWPGPSPESEFEGFTTH	300
Db	241	SLTTPSLDLPLTLTSLVILVLLTVLALLSHRRALKQKIWPGPSPESEFEGFTTH	300
Qy	301	KGNFQWLWYQNDGLWMSCTPTEDPPASLEVLSEKWCMTQAVRGCTDDEGFLLEPVG	360
Db	301	KGNFQWLWYQNDGLWMSCTPTEDPPASLEVLSEKWCMTQAVRGCTDDEGFLLEPVG	360
Qy	361	SEHAQDTYLVLDKWLPRNPSPEDLPQGGSDIVAMDEGSEASSCSALAKSPPEGAS	420
Db	361	SEHAQDTYLVLDKWLPRNPSPEDLPQGGSDIVAMDEGSEASSCSALAKSPPEGAS	420
Qy	421	AASFEYTLDPSSQLRPTWLCPELPPTPHLKYLVLVWSDSGISTDYSGDSQGAQGL	480
Db	388	-----PWALCELPPTPHLKYLVLVWSDSGISTDYSGDSQGAQGL	430
Qy	481	SDGPYSNPYENSILPAEPLPPSYVACS	508
Db	431	SDGPYSNPYENSILPAEPLPPSYVACS	458

RESULT 13

AD28671
 ID ADE28671 standard; protein; 458 AA.

XX ADE28671;

AC ADE28671;

DT 29-JAN-2004 (first entry)

DE Human NOV15a protein - SEQ ID 48.

XX NOVX; antidiabetic; anorectic; cardiast; hypotensive;

XX KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;

XX KW neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;

XX KW osteopathic; antiarthritic; antiinflammatory; dermatological;

KW antiaesthatic; antilipaemic; metabolic; diabetes; obesity; infectious;
 KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
 KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
 KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
 KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
 KW wound healing; angiogenesis; gene therapy; chromosome mapping;
 KW tissue typing; human; NOV.
 XX Homo sapiens.
 OS
 XX WO2003040330-A2.
 PN
 XX 15-MAY-2003.
 .PD
 XX
 XX 05-NOV-2002; 2002WO-US035536.
 XX
 XX 05-NOV-2001; 2001US-0338626P.
 PR 05-DEC-2001; 2001US-0336600P.
 PR 07-DEC-2001; 2001US-0338285P.
 PR 12-DEC-2001; 2001US-0341346P.
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 20-DEC-2001; 2001US-0342592P.
 PR 27-DEC-2001; 2001US-0344297P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383534P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0383829P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 26-AUG-2002; 2002US-0406353P.
 PR 31-OCT-2002; 2002US-00287971.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 XX Alsobrook JP, Alvarez E, Anderson DM, Baron M, Boldog FL;
 PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
 PI Ellerman K, Rittenberg S, Gangolli EA, Gerlach VL, Gorman L;
 PI Grose WM, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
 PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;
 PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
 PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
 PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
 PI Vernet CAM, Zerhusen BD, Zhong M;
 XX
 XX WPI; 2003-441555/41.
 DR N-PSDB; ADE28670.
 DR
 XX
 XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 XX Claim 1; SEQ ID NO 48; 447pp; English.
 XX
 XX The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
 CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
 CC fungicide, protozoacide, neurotropic, neuroprotective, antiparkinsonian,
 CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
 CC dermatological, antiaesthatic and antilipaemic activities. The
 CC polypeptides, nucleic acid molecules and antibodies may be useful for
 CC treating or diagnosing diseases including metabolic disorders such as
 CC diabetes and obesity, infectious diseases, anorexia, cancer,
 CC cardiovascular diseases including hypertension and atherosclerosis,
 CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
 CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
 CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.
 CC Furthermore, the nucleic acids and polypeptides may also be used to
 CC identify molecules that modulate or inhibit neurogenesis, cell

CC differentiation and proliferation, haemopoiesis, wound healing and
CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
CC be used as hybridisation probes, in chromosome mapping, tissue typing,
CC preventive medicine and pharmacogenomics. The current sequence is that of
CC the human NOV protein of the invention.
XX
SQ Sequence 458 AA;

Query Match 88.5%; Score 2404; DB 7; Length 458;
Best Local Similarity 89.2%; Pred. No. 1.1e-190;
Matches 453; Conservative 1; Mismatches 4; Indels 50; Gaps 1;
QY 1 MDHLGASLPQVGSGLLLAGAAWAPPNLPKPKESKAALLAARGPELLCFTTERLDL 60
Db 1 MDHLGASLPQVGSGLLLAGAAWAPPNLPKPKESKAALLAARGPELLCFTTERLDL 60
QY 61 VCFWEEAASAGVPGNYSFYQLEDEPKLCLRHQAAPTARGAVRWCSLPTADTSSFVPL 120
Db 61 VCFWEEAASAGVPGNYSFYQLEDEPKLCLRHQAAPTARGAVRWCSLPTADTSSFVPL 120
QY 121 ELRVTAASGAPRYHRVHINEVLLDAPVGLVARLADESGHVVLRLWLPPEPMTSHIRY 180
Db 121 ELRVTAASGAPRYHRVHINEVLLDAPVGLVARLADESGHVVLRLWLPPEPMTSHIRY 180
QY 181 EVDVSAGNAGSVQVREILLEGTECVLSNLCGRTRYTAVRARMAEPFGGFWSAWSEPV 240
Db 181 EVDVSAGNAGSVQVREILLEGTECVLSNLCGRTRYTAVRARMAEPFGGFWSAWSEPV 240
QY 241 SLTFTSDLDPLTLTSLILVILVLLTVALLSHRRALKQKIWPGIPSESEFEGFLFTTH 300
Db 241 SLTFTSDLDPLTLTSLILVILVLLTVALLSHRRALKQKIWPGIPSESEFEGFLFTTH 300
QY 301 KGNFQWLQYNDGCLWNSPCTPFTDDPASLEVLSERCWGTQVAPEPTDDEGPLEPVG 360
Db 301 KGNFQWLQYNDGCLWNSPCTPFTDDPASLEVLSERCWGTQVAPEPTDDEGPLEPVG 360
QY 361 SEHAQDTVLVLDKWLPRNPSEDLPGGGSDVIVADDEGSEASSCALASKSPGAS 420
Db 361 SEHAQDTVLVLDKWLPRNPSEDLPGGGSDVIVADDEGSEASSCALASKSPGAS 420
QY 421 AASFEYTLDPSSQLLRPWTLCPELPPTPHKLYLVVSDSGISTDYSSGDSQAQOGL 480
Db 388 -----FWALCPPELPPTPHKLYLVVSDSGISTDYSSGDSQAQOGL 430
QY 481 SDGYSNPNYENSLPAAEPLPPSYVACS 508
Db 431 SDGYSNPNYENSLPAAEPLPPSYVACS 458

RESULT 14
ADE28673 standard; protein; 458 AA.
XX
AC ADE28673;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human NOV15b protein - SEQ ID 50.
XX
KW NOVX; antidiabetic; anorectic; cardiatic; hypotensive;
KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
KW neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;
KW osteopathic; antiarthritic; antiinflammatory; dermatological;
KW antiaethmatic; antilipaeamic; metabolic; diabetes; obesity; infectious;
KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
KW wound healing; angiogenesis; gene therapy; chromosome mapping;
KW tissue typing; human; NOV.

Homo sapiens.

OS

XX

PN WO2003040330-A2.
XX
PD 15-MAY-2003.
XX
PF 05-NOV-2002; 2002WO-US035536.
XX
PR 05-NOV-2001; 2001US-0339626P.
PR 05-DEC-2001; 2001US-0336600P.
PR 07-DEC-2001; 2001US-0338285P.
PR 12-DEC-2001; 2001US-0341346P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 20-DEC-2001; 2001US-0342592P.
PR 27-DEC-2001; 2001US-0344297P.
PR 31-DEC-2001; 2001US-0344903P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 07-AUG-2002; 2002US-0401788P.
PR 26-AUG-2002; 2002US-0406353P.
PR 31-OCT-2002; 2002US-00287971.
XX
PA (CURA-) CURAGEN CORP.

Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FT,
Burgess CB, Caeman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
Ellerman K, Guo X, Hackett C, Ji W, Kekuda R, Khramtsov NV;
Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
Pena CEA, Reyman JA, Rastelli L, Rieger DK, Shenoy SG, Shimkets RA;
Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
Vernet CAM, Zerhusen BD, Zhong M;
WPI; 2003-441555/41.
N-PSDB; ADE28672.

New isolated NOVX polypeptides and polynucleotides, useful for
preventing, diagnosing or treating NOVX-associated disorders, e.g.
osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
asthma, or infections.

Claim 1; SEQ ID NO 50; 447pp; English.

The invention relates to a novel isolated NOVX polypeptide. The
polypeptide of the invention demonstrates, antidiabetic, anorectic,
cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
fungicide, protozoacide, neurotropic, neuroprotective, antiparkinsonian,
anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
dermatological, antiaethmatic and antilipaeamic activities. The
polypeptides, nucleic acid molecules and antibodies may be useful for
treating or diagnosing diseases including metabolic disorders such as
diabetes and obesity, infectious diseases, anorexia, cancer,
cardiovascular diseases including hypertension and atherosclerosis,
neurodegenerative disorders such as Alzheimer's disease, Parkinson's
disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
disorders, inflammatory skin disorders, asthma and dyslipidaemia.
Furthermore, the nucleic acids and polypeptides may also be used to
identify molecules that modulate or inhibit neurogenesis, cell
differentiation and proliferation, haemopoiesis, wound healing and
angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
be used as hybridisation probes, in chromosome mapping, tissue typing,
preventive medicine and pharmacogenomics. The current sequence is that of
the human NOV protein of the invention.

Sequence 458 AA;

Query Match 88.5%; Score 2404; DB 7; Length 458;

Best Local Similarity 89.2%; Pred. No. 1.1e-190;

Matches 453; Conservative 1; Mismatches 4; Indels 50; Gaps 1;
 QY 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFSKAAALAAARPELLCFTERLEDL 60
 DB 1 MDHLGASLWPGVSLCLLAGAAWAPPNLPDPKFSKAAALAAARPELLCFTERLEDL 60
 QY 61 VCFWEBAASAGVPGNYSYOLEDPWKLRLHQAPTARGAVRWCSLPTADTSSFVPL 120
 DB 61 VCFWEBAASAGVPGNYSYOLEDPWKLRLHQAPTARGAVRWCSLPTADTSSFVPL 120
 QY 121 ELRVTAASGAPRHRVHINEVLLDAPVGLVARLADESGHVLRWLPPEPMTSHIRY 180
 DB 121 ELRVTAASGAPRHRVHINEVLLDAPVGLVARLADESGHVLRWLPPEPMTSHIRY 180
 QY 181 EYDVSAGNAGSVQRVEILLEGTECVLSNLRGRTRYTFAVRARMAEPSPFGGFWSAWSEPV 240
 DB 181 EYDVSAGNAGSVQRVEILLEGTECVLSNLRGRTRYTFAVRARMAEPSPFGGFWSAWSEPV 240
 QY 241 SLLTSPDLDPLILTLVLLVILVLLTVLALLSHRRALKQKIWPGPSPESEFGLFTTH 300
 DB 241 SLLTSPDLDPLILTLVLLVILVLLTVLALLSHRRALKQKIWPGPSPESEFGLFTTH 300
 QY 301 KGNFOLWLYONDGCLWWSCTPFTEDPPASLEVLSEFCWGTQOAVEPGCTDDGGLLEPVG 360
 DB 301 KGNFOLWLYONDGCLWWSCTPFTEDPPAFLEVLSEFCWGTQOAVEPGCTDDGGLLEPVG 360
 QY 361 SEHAQDTYLVLDKWLPRNPSPSEDLPFGCGSVDIVAMDEGSEASSCSALASKPSPEGAS 420
 DB 361 SEHAQDTYLVLDKWLPRNPSPSEDLPFG----- 387
 QY 421 AASFEYTLDPSSQLLRPWLCPPELPPPTPHKYLVLVWSDSGISTDYSSGDSQAQOGL 480
 DB 388 -----PWALCPPELPPPTPHKYLVLVWSDSGISTDYSSGDSQAQOGL 430
 QY 481 SDGPYSNPVENSILIPAAEPLPSYVACS 508
 DB 431 SDGPYSNPVENSILIPAAEPLPSYVACS 459

RESULT 15
 ADM93416
 ID ADM93416 standard; protein; 458 AA.
 AC ADM93416;
 DT 01-JUL-2004 (first entry)
 DE Human NOVX polypeptide #24.
 XX
 KW Human; NOVX; congenital heart defect; cardiomyopathy; atherosclerosis;
 KW hypertension; pulmonary stenosis; scleroderma; adenocarcinoma;
 KW haemophilia; graft-versus-host disease; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW multiple sclerosis; diabetes; obesity; bronchial asthma;
 KW acquired immunodeficiency syndrome; AIDS; Crohn's disease;
 KW infectious disease; anorexia; immune disorder.
 XX
 OS Homo sapiens.
 XX
 PN US2004067882-A1.
 XX
 PD 08-APR-2004.
 XX
 PF 05-NOV-2002; 2002US-00287971.
 XX
 PR 22-OCT-2001; 2001US-00035568.
 PR 05-NOV-2001; 2001US-0338626P.
 PR 06-NOV-2001; 2001US-0333072P.
 PR 09-NOV-2001; 2001US-0345398P.
 PR 09-NOV-2001; 2001US-0348283P.
 PR 15-NOV-2001; 2001US-0335610P.
 PR 21-NOV-2001; 2001US-0332152P.
 PR 28-NOV-2001; 2001US-0333912P.

PR 29-NOV-2001; 2001US-00997425.
 PR 29-NOV-2001; 2001US-0334300P.
 PR 04-DEC-2001; 2001US-0336576P.
 PR 05-FEB-2002; 2002US-0354807P.
 PR 15-MAY-2002; 2002US-0380968P.
 PR 16-MAY-2002; 2002US-0381043P.
 PR 02-JUL-2002; 2002US-0393148P.
 PR 02-JUL-2002; 2002US-0393262P.
 PR 06-AUG-2002; 2002US-0401479P.
 PR 06-AUG-2002; 2002US-0401626P.
 PR 07-AUG-2002; 2002US-0401593P.
 PR 07-AUG-2002; 2002US-0401695P.
 PR 26-AUG-2002; 2002US-0406181P.
 XX
 PA (ALSO/) ALSOBROOK J P.
 PA (ALVA/) ALVAREZ E.
 PA (ANDE/) ANDERSON D W.
 PA (BARO/) BARON M.
 PA (BOLD/) BOLDOG F L.
 PA (BURG/) BURGESS C E.
 PA (CASM/) CASMAN S J.
 PA (CHAP/) CHAPOVAL A.
 PA (DHAN/) DHANABAL M.
 PA (EDIN/) EDINGER S R.
 PA (EISE/) EISEN A.
 PA (ELLE/) ELLERMAN K.
 PA (ETTE/) ETTERBERG S.
 PA (GANG/) GANGOLLI E A.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (GROS/) GROSSE W M.
 PA (GUOX/) GUO X.
 PA (HACK/) HACKETT C.
 PA (JIWV/) JI W.
 PA (KEKU/) KEKUDA R.
 PA (KIRA/) KHRANTSOV N V.
 PA (LEPL/) LEPLEY D M.
 PA (LILL/) LI L.
 PA (MACD/) MACDOUGALL J R.
 PA (MALY/) MALYANKAR U M.
 PA (MAZU/) MAZUR A.
 PA (MCQU/) MCQUEENEY K.
 PA (MEZE/) MEZES P S.
 PA (MILL/) MILLER C E.
 PA (MISH/) MISHRA V.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENNA C E A.
 PA (PEYM/) PEYMAN J A.
 PA (RAST/) RASTELLI L.
 PA (RIEG/) RIEGER D K.
 PA (ROTH/) ROTHENBERG M E.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SMIT/) SMITHSON G.
 PA (SPAD/) SPADERNA S K.
 PA (STAR/) STARLING G.
 PA (SPYT/) SPYTEK K A.
 PA (STON/) STONE D J.
 PA (TCHE/) TCHERNEV V T.
 PA (TWOM/) TWOMLOW N.
 PA (VERN/) VERNET C A M.
 PA (ZERH/) ZERHUSEN B D.
 PA (VOSS/) VOSS E Z.
 PA (ZHON/) ZHONG M.
 XX
 PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
 PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
 PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman LJ;
 PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
 PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeney K;
 PI Mezes PS, Miller CE, Milet I, Mishra V, Padigar M, Patturajan M;

Search completed: April 21, 2005, 08:51:50
Job time : 191 secs

PI Pena CEA, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME;
PI Shenoy SG, Shinkets RA, Smithson G, Spaderna SK, Starling G;
PI Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM;
PI Zerhusen BD, Voss EZ, Zhong M;
XX WPI; 2004-355303/33.
DR N-PSDB; ADM93415.
XX
XX Novel isolated NOVX polypeptide useful treating or preventing disorders
PT or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
PT sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
XX
XX Claim 2; SEQ ID NO 48; 330pp; English.
XX
XX The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The NOVX polypeptides and polynucleotides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of the sequences in a mammalian subject, and for
CC treating or preventing a pathology associated with NOVX. The
CC polypeptides, polynucleotides and antibodies that bind immunospecifically
CC to the polypeptides are useful for treating or preventing disorders or
CC syndromes such as congenital heart defects, cardiomyopathy,
CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
CC immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
CC anorexia and immune disorders. This sequence represents a human NOVX
CC polypeptide of the invention. Note: The sequence data for this patent is
CC also available from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 458 AA;

Query Match 88.5%; Score 2404; DB 8; Length 458;
Best Local Similarity 89.2%; Pred No. 1.1e-190;
Matches 453; Conservative 1; Mismatches 4; Indels 50; Gaps 1;
QY 1 MDHLGASLWPGVSGSLCILLAGAAPPNLPDPKFSKAAALLAARGPEELLCTERLEDL 60
DB 1 MDHLGASLWPGVSGSLCILLAGAAPPNLPDPKFSKAAALLAARGPEELLCTERLEDL 60
QY 61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120
DB 61 VCFWEEAASAGVGPGNYSFSYQLEDEPWKLCRLHQAPTARGAVRWCSLPTADTSSFPVL 120
QY 121 ELRVTAASGAPRYHRVHIHNEVLLDAPVGLVARLADSGHVLRWLPPTPTMTSHIRY 180
DB 121 ELRVTAASGAPRYHRVHIHNEVLLDAPVGLVARLADSGHVLRWLPPTPTMTSHIRY 180
QY 181 EVDVSAGNGAGSVORVEILEGRTECVLSNLGRTRTYTFAVRARMAEPSPGPGFWSAWSEPV 240
DB 181 EVDVSAGNGAGSVORVEILEGRTECVLSNLGRTRTYTFAVRARMAEPSPGPGFWSAWSEPV 240
QY 241 SLITPSDLPLIILTLISLILVILVLLVLLSHRRALKOKIWPGPSPESEFGLFTTH 300
DB 241 SLITPSDLPLIILTLISLILVILVLLVLLSHRRALKOKIWPGPSPESEFGLFTTH 300
QY 301 KGNFQWLWYQNDGCLMWSPCTPTTDPDPAFLVLSERCMGTQAVPEGTDEGFLLEPVG 360
DB 301 KGNFQWLWYQNDGCLMWSPCTPTTDPDPAFLVLSERCMGTQAVPEGTDEGFLLEPVG 360
QY 361 SEHAQDTYLVLDKWLPRNPSPEDLPFGGSDVIDAMDEGSEASSCSSALASKPSPEGAS 420
DB 361 SEHAQDTYLVLDKWLPRNPSPEDLPFGGSDVIDAMDEGSEASSCSSALASKPSPEGAS 420
QY 421 AASFETIILDPSSQLLRPWTLCPELPPTPHLKXLYLVWSDSGISTDYSSGDSQGAQGL 480
DB 388 -----PWALCPPELPPTPHLKXLYLVWSDSGISTDYSSGDSQGAQGL 430
QY 481 SDGFPYNSPYENSLIPAAEPLPPSYVACS 508
DB 431 SDGFPYNSPYENSLIPAAEPLPPSYVACS 458

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 21:00:26 ; Search time 5262 Seconds
(without alignments)
11046.017 Million cell updates/sec

Title: US-09-016-159D-4
Perfect score: 1527
Sequence: 1 atggaccactcggggcgctc.....gctatgtgcttctcttag 1527

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	1759	3	CR592865 full-leng
2	1527	100.0	1797	3	CR620075 full-leng
3	1527	100.0	1811	3	CR618473 full-leng
4	1455	95.3	1527	9	AY414846 Homo sapi
5	1413	92.5	1870	3	CR610088 full-leng
6	1271.6	83.3	1681	3	CR613702 full-leng
7	1005.8	65.9	1452	9	AY414848 Mus muscu
8	977.6	64.0	1082	5	BX382170 BX382170
9	949.6	62.2	1524	3	AK010968 Mus muscu
10	943	61.8	1320	3	CR597800 full-leng
11	863.8	56.6	984	5	BX448344 BX448344
12	857.4	56.1	1079	1	AL553257 AL553257
13	812.2	53.2	1005	5	BX335579 BX335579
14	792.2	51.9	1051	5	BX448343 BX448343
15	789	51.7	1115	9	AY414847 Pan trogl
16	773.4	50.6	775	5	BX096321 BX096321
17	762.4	49.9	968	5	BX406157 BX406157
18	749.8	49.1	997	1	AL523468 AL523468
19	747.4	48.9	1075	1	AL576311 AL576311
20	745.6	48.8	1127	1	AL523894 AL523894
21	722.4	47.3	918	5	BX378570 BX378570
22	693.8	45.4	881	1	AL521907 AL521907
23	687	45.0	769	4	BM020450 BM020450
24	678	44.4	972	1	AL523467 AL523467

25	672	44.0	1231	3	CR592320	CR592320 full-leng
26	669.4	43.8	994	5	BQ894763	BQ894763 AGENCOURT
27	653.4	42.8	1051	1	AL523895	AL523895 AGENCOURT
28	640	41.9	1030	5	BX340440	BX340440 BX340440
29	634.2	41.5	640	4	BI760304	BI760304 603045374
30	630	41.3	902	1	AL523633	AL523633 AL523633
31	610.4	40.0	816	5	BX406158	BX406158 BX406158
32	604.4	39.6	973	5	BX335578	BX335578 BX335578
33	561.8	36.8	945	1	AL523632	AL523632 AL523632
34	559.4	36.6	944	1	AL521906	AL521906 AL521906
35	558.4	36.6	1009	6	BY710355	BY710355 BY710355
36	544	35.6	1098	5	BX340439	BX340439 BX340439
37	541.6	35.5	1054	5	BX382171	BX382171 BX382171
38	509.6	33.4	890	4	BG105276	BG105276 602312859
39	501.6	32.8	951	6	CD518752	CD518752 AGENCOURT
40	487.8	31.9	853	5	BQ19762	BQ19762 AGENCOURT
41	478.4	31.3	724	5	BQ773300	BQ773300 UI-H-P80-
42	470	30.8	488	6	CB269218	CB269218 1008125 H
43	449.2	29.4	1006	4	BI838739	BI838739 603082183
44	440	28.8	984	4	BG163892	BG163892 602343252
45	419	27.4	534	2	BF078672	BF078672 229067 MA

ALIGNMENTS

RESULT 1
CR592865
LOCUS
DEFINITION
full-length cDNA clone CS0D1073YP22 of Placentia Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR592865
VERSION
CR592865.1 GI:50473672
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1759)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1759)
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1759
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1073YP22"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 100.0%; Score 1527; DB 3; Length 1759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACCACTCGGGCGCTCTGGCCCCAGGTCGCTCTTGTCTCTCTCGCTCGCT 60
DB 44 ATGGACCACTCGGGCGCTCTCTGGCCCCAGGTCGCTCTCTCTCTCTCTCTCTCT 103

QY 61 GGGGGCGCTGGGCGCCCCCGCTTAACCTCCGGACCCCAAGTTTCAGAGCAAAAGCGGCC 120
 DB |||||
 QY 104 GGGGGCGCTGGGCGCCCCCGCTTAACCTCCGGACCCCAAGTTTCAGAGCAAAAGCGGCC 163
 DB |||||
 QY 121 TTGCTGGGGCGGGGGCGGAGAGCTTCTGTGCTTCAACGAGCGGTGGAGACTTG 180
 DB |||||
 QY 164 TTGCTGGGGCGGGGGCGGAGAGCTTCTGTGCTTCAACGAGCGGTGGAGACTTG 223
 DB |||||
 QY 181 GTGTGTTTCTGGGAGGAGCGCGCTGGGGTGGCGCGGGCAACTACAGCTTCTCC 240
 DB |||||
 QY 224 GTGTGTTTCTGGGAGGAGCGCGCTGGGGTGGCGCGGGCAACTACAGCTTCTCC 283
 DB |||||
 QY 241 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCTTGCACACAGGCTCCACGCGCTCGT 300
 DB |||||
 QY 284 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCTTGCACACAGGCTCCACGCGCTCGT 343
 DB |||||
 QY 301 GGTGGGTGCGCTTCTGTGTTGCTGCTACAGCCGACAGTCCAGGCTTGTGCGCCCTA 360
 DB |||||
 QY 344 GGTGGGTGCGCTTCTGTGTTGCTGCTACAGCCGACAGTCCAGGCTTGTGCGCCCTA 403
 DB |||||
 QY 361 GAGTTGCGGCTCACAGAGCGCTCCGGCGCTCCGGGATATCACGCTGTATCCACATCAAT 420
 DB |||||
 QY 404 GAGTTGCGGCTCACAGAGCGCTCCGGCGCTCCGGGATATCACGCTGTATCCACATCAAT 463
 DB |||||
 QY 421 GAAATGAGTCTCTAGACGCGCCCGTGGGGTGTGTGGCGCGGTGGCTGACGAGCGGC 480
 DB |||||
 QY 464 GAAATGAGTCTCTAGACGCGCCCGTGGGGTGTGTGGCGCGGTGGCTGACGAGCGGC 523
 DB |||||
 QY 481 CACGTAGTGTGCGTGTGCTCCGGCGCTGAGACACCCATGACGTCTCACATCGGTAC 540
 DB |||||
 QY 524 CACGTAGTGTGCGTGTGCTCCGGCGCTGAGACACCCATGACGTCTCACATCGGTAC 583
 DB |||||
 QY 541 GAGTGGACGTCTCGGCGCGCAACCGCGCAGGAGCGTACAGAGGTGAGATCCTGGAG 600
 DB |||||
 QY 584 GAGTGGACGTCTCGGCGCGCAACCGCGCAGGAGCGTACAGAGGTGAGATCCTGGAG 643
 DB |||||
 QY 601 GGCGCACCGAGTGTCTGAGCAACCTGGGGGCGGAGCGCGTACACCTTCCGCGTC 660
 DB |||||
 QY 644 GGCGCACCGAGTGTCTGAGCAACCTGGGGGCGGAGCGCGTACACCTTCCGCGTC 703
 DB |||||
 QY 661 CGCGCGCTATGCTGAGCGAGCTTCCGGCGCTTCTGGAGCGCTGTGCGAGCGCTGTG 720
 DB |||||
 QY 704 CGCGCGCTATGCTGAGCGAGCTTCCGGCGCTTCTGGAGCGCTGTGCGAGCGCTGTG 763
 DB |||||
 QY 721 TCGTGTGAGCGCTAGCGACTTGAACCCCTCTACCTGAGCGCTCTCCCTCATCTCGTG 780
 DB |||||
 QY 764 TCGTGTGAGCGCTAGCGACTTGAACCCCTCTACCTGAGCGCTCTCCCTCATCTCGTG 823
 DB |||||
 QY 781 GTCATCTGTGTGCTGACCGTGTGCTGCGCTGTCTCCCAACCGCGGGCTCTGAAGCAG 840
 DB |||||
 QY 824 GTCATCTGTGTGCTGACCGTGTGCTGCGCTGTCTCCCAACCGCGGGCTCTGAAGCAG 883
 DB |||||
 QY 841 AAGATCTGGCTGATCCGAGCCGACAGAGCGAGTTTGAAGGCTTTCACCAACCCAC 900
 DB |||||
 QY 884 AAGATCTGGCTGATCCGAGCCGACAGAGCGAGTTTGAAGGCTTTCACCAACCCAC 943
 DB |||||
 QY 901 AAGGTAACTTCCAGCTGTGCTGTACAGAAATGATGCTGCTGTGTGAGCGCCCTGCG 960
 DB |||||
 QY 944 AAGGTAACTTCCAGCTGTGCTGTACAGAAATGATGCTGCTGTGTGAGCGCCCTGCG 1003
 DB |||||
 QY 961 ACCCCCTTTCAGGAGGACCCACCTGCTTCTTCTGGAAGTCTCTCAGAGCGTGTGCGGG 1020
 DB |||||
 QY 1004 ACCCCCTTTCAGGAGGACCCACCTGCTTCTTCTGGAAGTCTCTCAGAGCGTGTGCGGG 1063
 DB |||||
 QY 1021 ACGATGAGGAGTGGAGCGGGGACAGATATGAGGCGCCCTGCTGAGCGCGTGTGCGG 1080
 DB |||||
 QY 1064 ACGATGAGGAGTGGAGCGGGGACAGATATGAGGCGCCCTGCTGAGCGCGTGTGCGG 1123
 DB |||||
 QY 1081 AGTGAGATGCCAGGATACCTATCTGTGTGCTGAGCAAAATGTTGCTGCGCCCGGAACCG 1140
 DB |||||
 QY 1124 AGTGAGATGCCAGGATACCTATCTGTGTGCTGAGCAAAATGTTGCTGCGCCCGGAACCG 1183
 DB |||||

QY 1141 CCAGTGAGGACCTCCAGGCGCTGGTGGCAGTGTGACATAGTGGCCATGGATGAAGGC 1200
 DB |||||
 QY 1184 CCAGTGAGGACCTCCAGGCGCTGGTGGCAGTGTGACATAGTGGCCATGGATGAAGGC 1243
 DB |||||
 QY 1201 TCAGAGCATCTCTGCTCATCTGCTTTGGCCTCGAAGCCAGCCAGGAGGAGCTCT 1260
 DB |||||
 QY 1244 TCAGAGCATCTCTGCTCATCTGCTTTGGCCTCGAAGCCAGCCAGGAGGAGCTCT 1303
 DB |||||
 QY 1261 GCTGCCAGCTTTGAGTACATCTCTCGGACCCAGCTCCAGCTCTTTCGCTCCATGGACA 1320
 DB |||||
 QY 1304 GCTGCCAGCTTTGAGTACATCTCTCGGACCCAGCTCCAGCTCTTTCGCTCCATGGACA 1363
 DB |||||
 QY 1321 CTGTGCGCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTGTGTATCT 1380
 DB |||||
 QY 1364 CTGTGCGCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTGTGTATCT 1423
 DB |||||
 QY 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGAGGACTCCAGGAGGAGCCCAAGGGGCTTA 1440
 DB |||||
 QY 1424 GACTCTGGCATCTCAACTGACTACAGCTCAGGAGGACTCCAGGAGGAGCCCAAGGGGCTTA 1483
 DB |||||
 QY 1441 TCCGATGGCCCTTACTCCAAACCTTATGAGAACAGCCTTATCCAGCGCTGAGCGCTCTG 1500
 DB |||||
 QY 1484 TCCGATGGCCCTTACTCCAAACCTTATGAGAACAGCCTTATCCAGCGCTGAGCGCTCTG 1543
 DB |||||
 QY 1501 CCCCCAGCTATGTGGCTTGTCTTAG 1527
 DB |||||
 QY 1544 CCCCCAGCTATGTGGCTTGTCTTAG 1570
 DB |||||

RESULT 2

CR620075 1797 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0D1016YL05 of Placenta Cot 25-normalized
 DEFINITION of Homo sapiens (human).

ACCESSION CR620075.1 GI:50500882

VERSION HNC; CNSLT_CDNA.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1797)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

REFERENCE 2 (bases 1 to 1797)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT - Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

FEATURES

source

1..1797

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1016YL05"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 100.0%; Score 1527; DB 3; Length 1797;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACCACTCGGGGCGTCCCTCTGGGCCCCAGGTCGGCTCCCTTTGTCTCTCTGCTCGCT 60
 DB 93 ATGGACCACTCGGGGCGTCCCTCTGGGCCCCAGGTCGGCTCCCTTTGTCTCTCTGCTCGCT 152
 QY 61 GGGGCGGCTGGGCGCCCCCGCTTAACCTCCCGGACCCCAAGCTTCAGAGCAAGGCGCC 120
 DB 153 GGGGCGGCTGGGCGCCCCCGCTTAACCTCCCGGACCCCAAGCTTCAGAGCAAGGCGCC 212
 QY 121 TTGCTGGGCGGCCCCGGGGCCCGAAGAGCTTCTGTGCTTCCAGGAGCGGTTGGAGACTTG 180
 DB 213 TTGCTGGGCGGCCCCGGGGCCCGAAGAGCTTCTGTGCTTCCAGGAGCGGTTGGAGACTTG 272
 QY 181 GGTGTTTCTGGAGGAAGCGGCGAGCGCTGGGGTGGGCGCGGCAATACAGCTTCTCC 240
 DB 273 GTGTGTTTCTGGAGGAAGCGGCGAGCGCTGGGGTGGGCGCGGCAATACAGCTTCTCC 332
 QY 241 TACAGCTCGAGGATGAGCATGGAAGCTGTGCTCCGCTGCACACAGCTCCACAGGCTCGT 300
 DB 333 TACAGCTCGAGGATGAGCATGGAAGCTGTGCTCCGCTGCACACAGCTCCACAGGCTCGT 392
 QY 301 GGTGCGGTGCGCTTCTGTGTTTCTGCTGCTACAGCGCGACACAGCTTCGAGCTTCTGTCGCCCTA 360
 DB 393 GGTGCGGTGCGCTTCTGTGTTTCTGCTGCTACAGCGCGACAGCTTCGAGCTTCTGTCGCCCTA 452
 QY 361 GAGTTGCGGCTCACAGAGCCTTCGGGCTCCGCGATATCACCGTGTCAATCAATCAAT 420
 DB 453 GAGTTGCGGCTCACAGAGCCTTCGGGCTCCGCGATATCACCGTGTCAATCAATCAAT 512
 QY 421 GAAGTAGTCTCTAGAGCGCCCCGCTGGGCTGTGGCGCGGTGGCTGACGAGCGCG 480
 DB 513 GAAGTAGTCTCTAGAGCGCCCCGCTGGGCTGTGGCGCGGTGGCTGACGAGCGCG 572
 QY 481 CACGTAGTGTGCTGCTGCTGCTCCCGCGCTGAGACACCATGACGCTCTCACATCCGCTAC 540
 DB 573 CACGTAGTGTGCTGCTGCTGCTCCCGCGCTGAGACACCATGACGCTCTCACATCCGCTAC 632
 QY 541 GAGTGGAGCGTCTCGGCGCGCAACCGCGAGGAGCGGTACAGAGGAGGAGATCTCGAG 600
 DB 633 GAGTGGAGCGTCTCGGCGCGCAACCGCGAGGAGCGGTACAGAGGAGGAGATCTCGAG 692
 QY 601 GGCGCCAGCGAGTGTCTGAGCAACTGCGGGCGCGAGCGCTACACCTTCGCGCTC 660
 DB 693 GGCGCCAGCGAGTGTCTGAGCAACTGCGGGCGCGAGCGCTACACCTTCGCGCTC 752
 QY 661 CGCGCGCTATGCTGAGCGGAGCTTCGCGCGCTTCTGAGCGCTGCTCGAGGAGCTGTG 720
 DB 753 CGCGCGCTATGCTGAGCGGAGCTTCGCGCGCTTCTGAGCGCTGCTCGAGGAGCTGTG 812
 QY 721 TCGTGTGAGCGCTTACGAGCTGAGCGGCTTACGCGCTTCTGAGCGCTTCTGAGCGCTG 780
 DB 813 TCGTGTGAGCGCTTACGAGCTGAGCGGCTTACGCGCTTCTGAGCGCTTCTGAGCGCTG 872
 QY 781 GTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 873 GTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 932
 QY 841 AAGATCTGCGCTGGCATCCCGAGCCAGAGCGGAGTTGAAGGCTCTTACACACCCAC 900
 DB 933 AAGATCTGCGCTGGCATCCCGAGCCAGAGCGGAGTTGAAGGCTCTTACACACCCAC 992
 QY 901 AAGGTAATCTTACGCTGTGCTGTACAGAAATGATGCTGCTGTGTGTGAGGCGCTTCG 960
 DB 993 AAGGTAATCTTACGCTGTGCTGTACAGAAATGATGCTGCTGTGTGTGAGGCGCTTCG 1052
 QY 961 ACCCGCTTACGAGGAGCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGG 1020
 DB 1053 ACCCGCTTACGAGGAGCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGG 1112
 QY 1021 ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGAGCGCTGCTGGG 1080
 DB 1113 ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGAGCGCTGCTGGG 1172
 QY 1081 AGTGAGCATGCCAGGATACCTATCTGCTGTGGACAATGTTGCTGCCCGGAAACCCG 1140

DB 1173 AGTGAGCATGCCAGGATACCTATCTGCTGTGGACAATGTTGCTGCCCGGAACCCG 1232
 QY 1141 CCCAGTGAAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
 DB 1233 CCCAGTGAAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC 1292
 QY 1201 TCAGAAGCATCTCTCTGCTCATCTGCTTTGGCCCTCGAAGCCAGCCAGAGGAGGCTCT 1260
 DB 1293 TCAGAAGCATCTCTCTGCTCATCTGCTTTGGCCCTCGAAGCCAGCCAGAGGAGGCTCT 1352
 QY 1261 GCTGCCAGCTTTGAGTACACTATCTCGACCCAGCTCCAGCTCTTGGCTCCATGGACA 1320
 DB 1353 GCTGCCAGCTTTGAGTACACTATCTCGACCCAGCTCCAGCTCTTGGCTCCATGGACA 1412
 QY 1321 CTGTGCCCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTTGTGTATCT 1380
 DB 1413 CTGTGCCCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTTGTGTATCT 1472
 QY 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGAGGCCAAAGGGGCTTA 1440
 DB 1473 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGAGGCCAAAGGGGCTTA 1532
 QY 1441 TCCGATGGCCCCCTACTCAACCTTATGAGNACAGCTTATCCAGCGCTGAGCCTCTG 1500
 DB 1533 TCCGATGGCCCCCTACTCAACCTTATGAGNACAGCTTATCCAGCGCTGAGCCTCTG 1592
 QY 1501 CCCCCAGCTATGTGGCTTGTCTTAG 1527
 DB 1593 CCCCCAGCTATGTGGCTTGTCTTAG 1619

RESULT 3
 CR618473
 LOCUS 1811 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CSODM009Y22 of Fetal liver of Homo sapiens (human).
 ACCESSION CR618473
 VERSION CR618473.1 GI:50499280
 KEYWORDS HTC; CNSLT CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1811)
 AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1811)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 source
 1. 1811
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODM009Y22"
 /tissue type="Fetal liver"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Query Match 100.0%; Score 1527; DB 3; Length 1811;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGGACCACTCGGGCGTCCCTCTGGCCCAAGGTGGCTCCCTTGTCTCTCTCGTCCGT	60
DB	79	ATGGACCACTCGGGCGTCCCTCTGGCCCAAGGTGGCTCCCTTGTCTCTCTCGTCCGT	138
QY	61	GGGGCGCTGGGGCGCCCGGCTAACTCCCGGACCCCAAGTTCGAGAGCAAGCGGCC	120
DB	139	GGGGCGCTGGGGCGCCCGGCTAACTCCCGGACCCCAAGTTCGAGAGCAAGCGGCC	198
QY	121	TTGCTGGCGCGCGGGGCGGAGAGCTTCTGCTTCAACGAGGGTGGAGACTTG	180
DB	199	TTGCTGGCGCGCGGGGCGGAGAGCTTCTGCTTCAACGAGGGTGGAGACTTG	258
QY	181	GTGTGTTTCTGGGAGGAAGCGGCGAGCGTGGGGTGGGCGGCAACTACAGCTTCTCC	240
DB	259	GTGTGTTTCTGGGAGGAAGCGGCGAGCGTGGGGTGGGCGGCAACTACAGCTTCTCC	318
QY	241	TACAGCTCGAGATGAGCCATGGAAGCTGTGTGCTTGCAACAGGCTCCCAAGCGCTG	300
DB	319	TACAGCTCGAGATGAGCCATGGAAGCTGTGTGCTTGCAACAGGCTCCCAAGCGCTG	378
QY	301	GGTGGGTGGCTTCTGGTGTCTGCTGCTACAGCGGACAGCTGCTGCTGCGCCTA	360
DB	379	GGTGGGTGGCTTCTGGTGTCTGCTGCTACAGCGGACAGCTGCTGCTGCGCCTA	438
QY	361	GAGTTGGCGGTCAACAGCAGCTCCGGCGCTCCGGCATATCACCGTGTATCACATCAAT	420
DB	439	GAGTTGGCGGTCAACAGCAGCTCCGGCGCTCCGGCATATCACCGTGTATCACATCAAT	498
QY	421	GAAGTAGTGTCTCTAGACGCCCCCGTGGGGCTGGTGGCGGCTTGGCTGACGAGAGCGC	480
DB	499	GAAGTAGTGTCTCTAGACGCCCCCGTGGGGCTGGTGGCGGCTTGGCTGACGAGAGCGC	558
QY	481	CAGTATGTTGGCTGGCTCCGGCGGCTGAGACACCAATGACGCTCTCACATCCGCTAC	540
DB	559	CAGTATGTTGGCTGGCTCCGGCGGCTGAGACACCAATGACGCTCTCACATCCGCTAC	618
QY	541	GAGTGGAGCTCTCGCGCGGCAACCGCGCAGGAGCGTACAGAGGTTGAGATCTCGGAG	600
DB	619	GAGTGGAGCTCTCGCGCGGCAACCGCGCAGGAGCGTACAGAGGTTGAGATCTCGGAG	678
QY	601	GGCCGACCGAGTGTGTGTAGAACCTGTGGGGCGGACCGGCTACACCTTCCGCTC	660
DB	679	GGCCGACCGAGTGTGTGTAGAACCTGTGGGGCGGACCGGCTACACCTTCCGCTC	738
QY	661	CGCGCGGTATGGCTGAGCCGAGCTTCCGGCGCTTCTGGAGCGCTGTGGAGGCTGTG	720
DB	739	CGCGCGGTATGGCTGAGCCGAGCTTCCGGCGCTTCTGGAGCGCTGTGGAGGCTGTG	798
QY	721	TGCTGTGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT	780
DB	799	TGCTGTGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT	858
QY	781	GTATCTCTGGT	840
DB	859	GTATCTCTGGT	918
QY	841	AGATCTGGCTGGATCCCGAGCCGAGAGCGAGTTTGAAGGCTCTTACACACCCAC	900
DB	919	AGATCTGGCTGGATCCCGAGCCGAGAGCGAGTTTGAAGGCTCTTACACACCCAC	978
QY	901	AAGGTAATCTTCAGCTGTGGCTGTACAGAAATGATGCTGTGTGTGTGTGTGTGTGTGT	960
DB	979	AAGGTAATCTTCAGCTGTGGCTGTACAGAAATGATGCTGTGTGTGTGTGTGTGTGTGT	1038
QY	961	ACCCCTTTCAGGAGGACCACTCTCTTCCCTGGAGTCTCTCTAGAGCGTGTGTGTGTGT	1020
DB	1039	ACCCCTTTCAGGAGGACCACTCTCTTCCCTGGAGTCTCTCTAGAGCGTGTGTGTGTGT	1098
QY	1021	ACGATGACGAGTGTGGAGCGGGGACAGATGATGAGGGCCCCCTCTGCTGAGCCAGTGGC	1080
DB	1099	ACGATGACGAGTGTGGAGCGGGGACAGATGATGAGGGCCCCCTCTGCTGAGCCAGTGGC	1158
QY	1081	AGTGAGCATGCCAGGATACCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1140
DB	1159	AGTGAGCATGCCAGGATACCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1218
QY	1141	CCGAGTGAAGACCTTCCAGGGCCCTGGTGGCAGTGTGTGTGTGTGTGTGTGTGTGTGT	1200
DB	1219	CCGAGTGAAGACCTTCCAGGGCCCTGGTGGCAGTGTGTGTGTGTGTGTGTGTGTGTGT	1278
QY	1201	TCAGAAGCATCTCTCTGCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1260
DB	1279	TCAGAAGCATCTCTCTGCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1338
QY	1261	GCTGCCAGCTTTGAGTACATCTCTTGGACCCAGCTTCCAGCTTCTTGCCTCCATGAGCA	1320
DB	1339	GCTGCCAGCTTTGAGTACATCTCTTGGACCCAGCTTCCAGCTTCTTGCCTCCATGAGCA	1398
QY	1321	CTGTGCCCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTGTGTGTATCT	1380
DB	1399	CTGTGCCCTGAGCTGCCCTTACCCACCCACCTAAAGTACCTGTACCTTGTGTGTATCT	1458
QY	1381	GACTTGTGCATCTCACTGACTACAGCTCAGGGGACTCCCAAGGAGGCCCAAGGGGGCTTA	1440
DB	1459	GACTTGTGCATCTCACTGACTACAGCTCAGGGGACTCCCAAGGAGGCCCAAGGGGGCTTA	1518
QY	1441	TCCGATGGCCCCCTACTCCAAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG	1500
DB	1519	TCCGATGGCCCCCTACTCCAAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG	1578
QY	1501	CCCCCAGCTATGTGGCTTGTCTTTAG 1527	
DB	1579	CCCCCAGCTATGTGGCTTGTCTTTAG 1605	
RESULT 4			
AY414846			
LOCUS			
DEFINITION	Homo sapiens EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,		
ACCESSION	AY414846		
VERSION	AY414846.1 GI:39770805		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1527) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios Science 302 (5652), 1960-1963 (2003)		
AUTHORS	2 (bases 1 to 1527) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment.		
TITLE	Location/Qualifiers		
JOURNAL	1. .1527		
PUBMED	/organism="Homo sapiens"		
REFERENCE	/mol_type="genomic DNA"		
AUTHORS	/db_xref="taxon:9606"		
	<1_->1527		
	/gene="EPOR"		
TITLE	source		
JOURNAL	1. .1527		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1. .1527		
gene	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	<1_->1527		
	/gene="EPOR"		

ORIGIN		/locus_tag="HCM5360"	
Query Match		95.3%; Score 1455; DB 9; Length 1527;	
Best Local Similarity		95.3%; Pred. No. 0;	
Matches 1455; Conservative		0; Mismatches 72; Indels 0; Gaps 0;	
QY	1	ATGGACCACTCGGGCGTCCCTCTGGCCCCAGGTGGCTCCCTTTGTTCTCTGCTCGCT	60
DB	1	ATGGACCACTCGGGCGTCCCTCTGGCCCCAGGTGGCTCCCTTTGTTCTCTGCTCGCT	60
QY	61	GGGGCGCTTGGGCGCCCCGCTTAACCTCCCGGACCCCAAGTTCCAGAGCAAGGGGCC	120
DB	61	NN	120
QY	121	TTGCTGGCGGCGCGGGGCGCGAAGAGCTTCTGTGCTTCCACGAGCGGTTGGAGACTTG	180
DB	121	TTGCTGGCGGCGCGGGGCGCGAAGAGCTTCTGTGCTTCCACGAGCGGTTGGAGACTTG	180
QY	181	GTGTGTTTCTGGAGGAAGCGGCGAGCGTGGGGTGGGCCCGGGCAACTACAGCTTCTCC	240
DB	181	GTGTGTTTCTGGAGGAAGCGGCGAGCGTGGGGTGGGCCCGGGCAACTACAGCTTCTCC	240
QY	241	TACCACTTCGAGATGAGCCATGGAAGCTGTGTGCTGTCACAGGCTCCGACGGCTCGT	300
DB	241	TACCACTTCGANNNNNNNNNNNNAAGCTGTGTGCTGTCACAGGCTCCGACGGCTCGT	300
QY	301	GGTGGGTCGGCTTCTGGTGTTCGCTGCTACAGCGGACACCTCGAGCTTCTGGGCCCTA	360
DB	301	GGTGGGTCGGCTTCTGGTGTTCGCTGCTACAGCGGACACCTCGAGCTTCTGGGCCCTA	360
QY	361	GAGTTGCGGCTCACAGCAGCTCCGGCGTCCGGCGATATCACCGTGTATCCACATCAAT	420
DB	361	GAGTTGCGGCTCACAGCAGCTCCGGCGTCCGGCGATATCACCGTGTATCCACATCAAT	420
QY	421	GAACTAGTCTCTAGACGCCCCGCTGGGGCTGGTGGCGGCTTGGCTGACGAGCGGC	480
DB	421	GAACTAGTCTCTAGACGCCCCGCTGGGGCTGGTGGCGGCTTGGCTGACGAGCGGC	480
QY	481	CACGTAGTCTTGGCTGGCTCCCGCGCTGAGACACCCATGACGTCTCACATCCGCTAC	540
DB	481	CACGTAGTCTTGGCTGGCTCCCGCGCTGAGACACCCATGACGTCTCACATCCGCTAC	540
QY	541	GAGTGGAGCTCTCGGCGGCAACCGCGCAGGAGCGTACAGAGGTTGGAGATCTCGAG	600
DB	541	GAGTGGAGCTCTCGGCGGCAACCGCGCAGGAGCGTACAGAGGTTGGAGATCTCGAG	600
QY	601	GGCGGACCGAGTGTGTGAGCAACCTCGCGGGCGGACGCGCTACACCTTCGCGCTC	660
DB	601	GGCGGACCGAGTGTGTGAGCAACCTCGCGGGCGGACGCGCTACACCTTCGCGCTC	660
QY	661	CGCGCGGTATGGCTGAGCGAGCTTCGGCGCTTCTGGAGCGCTTGGTGGAGCTGTG	720
DB	661	CGCGCGGTATGGCTGAGCGAGCTTCGGCGCTTCTGGAGCGCTTGGTGGAGCTGTG	720
QY	721	TCGCTGCTGACGCTAGCAGCTGGACCCCTCATCTGACGCTTCCCTCATCTCTCGTG	780
DB	721	TCGCTGCTGACGCTAGCAGCTGGACCCCTCATCTGACGCTTCCCTCATCTCTCGTG	780
QY	781	GTTCATCTGGTGTGTGACCGTGTGCGGCTGTCTCCACGCGCGGCTCTGAAGCAG	840
DB	781	GTTCATCTGGTGTGTGACCGTGTGCGGCTGTCTCCACGCGCGGCTCTGAAGCAG	840
QY	841	RAGATCTGCGCTGGATCCCGAGCCAGAGCGGAGTTTGAAGGCTCTTACACACCCAC	900
DB	841	RAGATCTGCGCTGGATCCCGAGCCAGAGCGGAGTTTGAAGGCTCTTACACACCCAC	900
QY	901	AAGGTAACCTCCAGCTGTGGCTGTACAGAAATGATGCTGCTGTGGTGGAGCCCTGCG	960
DB	901	AAGGTAACCTCCAGCTGTGGCTGTACAGAAATGATGCTGCTGTGGTGGAGCCCTGCG	960
QY	961	ACCCCTTCACGGAGACCCCACTGCTTCTCCCTGGAGTCTCTTCAGAGCGCTGTGGGG	1020
DB	961	ACCCCTTCACGGAGACCCCACTGCTTCTCCCTGGAGTCTCTTCAGAGCGCTGTGGGG	1020

RESULT 5
CR610088
LOCUS full-length cDNA clone CS0DC003YF02 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR610088.1 GI:50490895
VERSION CR610088
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li W.B., Gruber C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
Genoscope.
2 (bases 1 to 1870)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. 1870
/organism="Homo sapiens"
/mol_type="mRNA"


```

/db_xref="taxon:9606"
/clone="CS0D0003YF02"
/tissue="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

```

ORIGIN

```

Query Match      92.5%; Score 1413; DB 3; Length 1870;
Best Local Similarity 93.6%; Pred. No. 3.6e-311;
Matches 1527; Conservative 0; Mismatches 0; Indels 104; Gaps 1;

QY 1  ATGACACACCTCGGGGGTCCCTCTGCGCCAGGTGGCTCCCTTGTCTCCTGCTCGCT 60
DB 44  ATGGACACCTCGGGGGTCCCTCTGCGCCAGGTGGCTCCCTTGTCTCCTGCTCGCT 103
QY 61  GGGGCGGCTCGGGGGTCCCTTAACCTCCGAGACCCCAAGTTCAGAGCAAGCGGCC 120
DB 104  GGGGCGGCTCGGGGGTCCCTTAACCTCCGAGACCCCAAGTTCAGAGCAAGCGGCC 163
QY 121  TTGCTGGCGGCTCGGGGGTCCCTTAACCTCCGAGACCCCAAGTTCAGAGCAAGCGGCC 180
DB 164  TTGCTGGCGGCTCGGGGGTCCCTTAACCTCCGAGACCCCAAGTTCAGAGCAAGCGGCC 223
QY 181  GTCTGTCTCTGGGAGGAGCGGCGAGCGTGGGTTGGGCGGCAACTACAGCTTCTCC 240
DB 224  GTCTGTCTCTGGGAGGAGCGGCGAGCGTGGGTTGGGCGGCAACTACAGCTTCTCC 283
QY 241  TACAGCTCGAGGATGAGGCATGGAAGCTGTGTGCGCTGCACAGAGCTCCACGGCTCGT 300
DB 284  TACAGCTCGAGGATGAGGCATGGAAGCTGTGTGCGCTGCACAGAGCTCCACGGCTCGT 343
QY 301  GGTGCGGTGCGCTTCTGCTGTTTCCGCTTACAGCGGACAGCTGAGCTTGTGCGCCCTA 360
DB 344  GGTGCGGTGCGCTTCTGCTGTTTCCGCTTACAGCGGACAGCTGAGCTTGTGCGCCCTA 403
QY 361  GAGTTGGGCTCAGAGGCTCGGGGCTCGCGGATACCGGCTGTCATCCACATCAAT 420
DB 404  GAGTTGGGCTCAGAGGCTCGGGGCTCGCGGATACCGGCTGTCATCCACATCAAT 463
QY 421  GAAGTAGTGTCTCTAGACGCGGCTGGGCTGGTGGCGGTTGGCTGACAGAGCGGC 480
DB 464  GAAGTAGTGTCTCTAGACGCGGCTGGGCTGGTGGCGGTTGGCTGACAGAGCGGC 523
QY 481  CAGTAGTGTGTGGCTTCCGCGGCTGAGACACCCATGACGCTCAGATCCGCTAC 540
DB 524  CAGTAGTGTGTGGCTTCCGCGGCTGAGACACCCATGACGCTCAGATCCGCTAC 583
QY 541  GAGTGGAAGCTCTCGCGGCGGACGCGGAGGAGGCTACAGAG-----584
DB 584  GAGTGGAAGCTCTCGCGGCGGACGCGGAGGAGGCTACAGAGGGGCGACCGTTTCTCTT 643
QY 585 -----584
DB 644  TCCCTGATTTGGCTCAGTTCCACGAGGCGGCGGCGGCGGCTATTTTGTCTGCTA 703
QY 585 -----GGTGAGATCTCTGAGGCGGCGGCGGCGGCGGCTATTTTGTCTGCTA 703
DB 704  CGCGTCCCTCGCGCTGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 763
QY 617  TGCTGAGCAACTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 676
DB 764  TGCTGAGCAACTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 823
QY 677  AGCGGAGCTTCGCGGCGGCTTCGAGGCGGCTGCTGCGGAGCTGCTGCTGCTGACGCT 736
DB 824  AGCGGAGCTTCGCGGCGGCTTCGAGGCGGCTGCTGCGGAGCTGCTGCTGCTGACGCT 883
QY 737  GCGACCTGAGACCCCTCATCTGAGCGCTTCCCTCATCTGCTGCTGCTGCTGCTGCT 796
DB 884  GCGACCTGAGACCCCTCATCTGAGCGCTTCCCTCATCTGCTGCTGCTGCTGCTGCT 943
QY 797  TGACCGTGTGCGGCTGCTTCCACGCGGCGGCTTGAAGCAGAGATCTGCGCTGGCA 856
DB 944  TGACCGTGTGCGGCTGCTTCCACGCGGCGGCTTGAAGCAGAGATCTGCGCTGGCA 1003

```

```

QY 857  TCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTCAACCAACCAAGGTTAACTTCCAGC 916
DB 1004  TCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTCAACCAACCAAGGTTAACTTCCAGC 1063
QY 917  TGTGCTGTACCCAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
DB 1064  TGTGCTGTGTACCCAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
QY 977  ACCACCTGCTTCCCTCGAAGTCTCTCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
DB 1124  ACCACCTGCTTCCCTCGAAGTCTCTCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1183
QY 1037  AGCCGGGACAGATGATGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
DB 1184  AGCCGGGACAGATGATGAGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
QY 1097  ATACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156
DB 1244  ATACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1303
QY 1157  CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
DB 1304  CAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1363
QY 1217  GCTCATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276
DB 1364  GCTCATCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
QY 1277  ACACCTATCTGAGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1336
DB 1424  ACACCTATCTGAGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1483
QY 1337  CCCCTACCCACCCACCTAAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396
DB 1484  CCCCTACCCACCCACCTAAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1543
QY 1397  CTGACTACAGCTCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1456
DB 1544  CTGACTACAGCTCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1603
QY 1457  CCAACCCCTTATGAGAACAGCTTATCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1516
DB 1604  CCAACCCCTTATGAGAACAGCTTATCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1663
QY 1517  CTTGCTCTTAG 1527
DB 1664  CTTGCTCTTAG 1674

RESULT 6
LOCUS CR613702
DEFINITION full-length cDNA clone CS0D1076YC13 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR613702
VERSION CR613702.1
KEYWORDS GI:50494509
SOURCE HTC; CNSLT cdna.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1681)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Paradise Avenue
Genoscope.
2 (bases 1 to 1681)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

```

BP 191 91006 EVERY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES

source

Location/Qualifiers
 1. 1681
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D10761C13"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 83.3%; Score 1271.6; DB 3; Length 1681;
 Best Local Similarity 99.7%; Pred. No. 5.9e-279;
 Matches 1274; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

245 AGCTCGAGGATGAGCCATGGAAGCTGTGCGCTTGACACAGGCTCCACCGCTCGTGGTG 304
 404 ATCGAGGGGATGACCATGAGAGCTGTGCGCTGACAGGCTCCACCGCTCGTGGTG 463
 305 CGGTGCGCTTCTGCTGTGCTGCTACAGCCGACAGCTCGAGCTTCTGCTCCCTTAGAGT 364
 464 CGGTGCGCTTCTGCTGTGCTGCTACAGCCGACAGCTCGAGCTTCTGCTCCCTTAGAGT 523
 365 TGGCGCTCAGAGAGCTTCCGGGCTCCGGATATACCGTGTTCATCCATCAATGAAG 424
 524 TGGCGCTCAGAGAGCTTCCGGGCTCCGGATATACCGTGTTCATCCATCAATGAAG 583
 425 TAGTGTCTCTAGAGCGCCCGCTGGGGCTGTGGCGGTGGCTGAGAGAGCGGCCACG 484
 584 TAGTGTCTCTAGAGCGCCCGCTGGGGCTGTGGCGGTGGCTGAGAGAGCGGCCACG 643
 485 TAGTGTGCTGCTGCTCCCGCGCTCGAGACACCCATGAGCTTCTCATCCGTACGAGG 544
 644 TAGTGTGCTGCTGCTCCCGCGCTCGAGACACCCATGAGCTTCTCATCCGTACGAGG 703
 545 TGGAGCTCTCGCGCGCAAGCGCGAGGAGCGTACAGAGGGTGGAGATCCTCGAGGGCC 604
 704 TGGAGCTCTCGCGCGCAAGCGCGAGGAGCGTACAGAGGGTGGAGATCCTCGAGGGCC 763
 605 GCACCGAGTGTGTGAGCAACTCGCGGGCGGAGCGCTACACCTTCGCGCTCCGCG 664
 764 GCACCGAGTGTGTGAGCAACTCGCGGGCGGAGCGCTACACCTTCGCGCTCCGCG 823
 665 CGCGTATGGCTGAGCGAGCTTGGCGGCTTCTGGAGCGCTTGGTGGAGCGCTGTGTGCG 724
 824 CGCGTATGGCTGAGCGAGCTTGGCGGCTTCTGGAGCGCTTGGTGGAGCGCTGTGTGCG 883
 725 TGCTGAGCTTAGGAGCTTGACCCCTCATCTGAGCGCTTCCCTCATCTCTGCTGTGCTCA 784
 884 TGCTGAGCTTAGGAGCTTGACCCCTCATCTGAGCGCTTCCCTCATCTCTGCTGTGCTCA 943
 785 TCCTGGTGTGCTGACCGCTGTGCGCTGTCTCTCCACCGCGGCTTCTGAAGCAGAAGA 844
 944 TCCTGGTGTGCTGACCGCTGTGCGCTGTCTCTCCACCGCGGCTTCTGAAGCAGAAGA 1003
 845 TCTGGCTTGGCTCCCGAGCCGAGAGCGAGTTTGAAGCGCTTCTTCCACCGCCCAAGG 904
 1004 TCTGGCTTGGCTCCCGAGCCGAGAGCGAGTTTGAAGCGCTTCTTCCACCGCCCAAGG 1063
 905 GTAACCTCAGCTGTGGCTGTACAGATGATGGCTGTGGTGGAGCGCTTGCACCC 964
 1064 GTAACCTCAGCTGTGGCTGTACAGATGATGGCTGTGGTGGAGCGCTTGCACCC 1123
 965 CCTTCAGGAGGACCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGGACGA 1024
 1124 CCTTCAGGAGGACCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGGACGA 1183

Qy 1025 TGCGAGCTGGAGCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGCGAGTG 1084
 Db 1184 TGCAGGCGAGTGGAGCGGGGACAGATGATGAGGGCCCCCTGCTGGAGCCAGTGGGCGAGTG 1243
 Qy 1085 AGCATGCCAGGATACCTATCTGCTGCTGACAAATGGTTGCTGCCCGGAAACCCGCCCA 1144
 Db 1244 AGCATGCCAGGATACCTATCTGCTGCTGACAAATGGTTGCTGCCCGGAAACCCGCCCA 1303
 Qy 1145 GTGAGGACCTCCAGGGCCCTGGTGGCAGTGTGACATAGTGGCCATGATGAAGGCTCAG 1204
 Db 1304 GTGAGGACCTCCAGGGCCCTGGTGGCAGTGTGACATAGTGGCCATGATGAAGGCTCAG 1363
 Qy 1205 AAGCATCTCTGCTCATCTGCTTTGGCCTCGAGGCCAGCCAGAGGGAGCCTCTGCTG 1264
 Db 1364 AAGCATCTCTGCTCATCTGCTTTGGCCTCGAGGCCAGCCAGAGGGAGCCTCTGCTG 1423
 Qy 1265 CCAGCTTTGAGTACACTATCTGGACCCCGAGCTCCAGAGCTTTGCGTCCATGACACTGT 1324
 Db 1424 CCAGCTTTGAGTACACTATCTGGACCCCGAGCTCCAGAGCTTTGCGTCCATGACACTGT 1483
 Qy 1325 GCCTGAGCTGCCCCCTACCCACCCACCTAAAGTACCTGTACCTTTGTGGTATCTGACT 1384
 Db 1484 GCCTGAGCTGCCCCCTACCCACCCACCTAAAGTACCTGTACCTTTGTGGTATCTGACT 1543
 Qy 1385 CTGGCATCTCAACTGACTAGAGCTAGGGAGCTCCAGAGGAGCCCAAGGGGCTTATCCG 1444
 Db 1544 CTGGCATCTCAACTGACTAGAGCTAGGGAGCTCCAGAGGAGCCCAAGGGGCTTATCCG 1603
 Qy 1445 ATGGCCCCCTACTCCACCCCTTATGAGAACAGCTTATCCAGCGCTTGAGCCCTCTGCCCC 1504
 Db 1604 ATGGCCCCCTACTCCACCCCTTATGAGAACAGCTTATCCAGCGCTTGAGCCCTCTGCCCC 1663
 Qy 1505 CCAGCTATGTGGCTTGTCT 1522
 Db 1664 CCAGCTATGTGGCTTGTCT 1681

RESULT 7

AY414848

LOCUS

Mus musculus EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
 Genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

AY414848 1452 bp DNA linear GSS 17-DEC-2003
 Mus musculus EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
 Genomic survey sequence.
 AY414848
 AY414848.1 GI:39770807
 GSS.
 Mus musculus (house mouse)
 Mus musculus
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1452)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 Location/Qualifiers
 1. 1452
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

```
gene      <1..>1452
/locus_tag="HCW5360"
ORIGIN
Query Match      65.9%; Score 1005.8; DB 9; Length 1452;
Best Local Similarity 81.5%; Pred. No. 2.2e-218;
Matches 1195; Conservative 0; Mismatches 257; Indels 15; Gaps 2;
Qy 61 GGGCGCGCTGGGCGCCCGCCCTAACCTCCCGACCCCAAGTTTCGAGAGCAAGCGCC 120
Db 1 GGGGAGCGCTGGGCACTTCAACCGAGCTCCCGGACCCCAAGTTTCGAGAGCAAGCGCC 60
Qy 121 TTGTTGGCGGCGCGGGGCGCCGAAGAGCTTCTGTGCTTTCACCGAGCGGTTGAGAGACTTG 180
Db 61 CTGCTGGCATCCCGGGGCTCCGAGAACTTCTGTGCTTCAACCAACGCTTGGAGACTTG 120
Qy 181 GTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGTGGCGCCCGGGCAACTACAGTTCTCC 240
Db 121 GTGTGTTTCTGGAGGAAGCGCGAGCTCCGGGATGGAC---TTCAACTACAGTTCTCA 177
Qy 241 TACCAGCTCGAGGATGAGCCATGAAGCTGTGCTACAGCCGACAGCTCGAGCTTCCACGCGCTCGT 300
Db 178 TACCAGCTCGAGGATGAGTACGAAAGTCACTGTAGCTTGCACCAAGGCTCCACCGCTCCG 237
Qy 301 GGTGCGGTGCGCTTCTGTGTGCTTACAGCCGACAGCTCGAGCTTCTGTGCGCCCTA 360
Db 238 GGCTCCGTCGCTTCTGTGTGCTTCACTGCCAAGCGGACATCGAGTTTGTGCGCGCTG 297
Qy 361 GAGTTGCGGCTACAGCAGCTTCCGGCGCTCCGGGATATACCGTGTATCCATCAAT 420
Db 298 GAGCTGAGGTTGAGCGGCGCTCGGTTCTCTCGCTATCACCGCATCATCCATA----- 352
Qy 421 GAAGTAGTGTCTTAGAGCGCCCGTGGGCTGTGTGGCGGTTGGCTGAGGAGCGGC 480
Db 353 -----TGCTCTGGAGCGCCCGCGGGCTGTGTGGCGCCCGGGCAGAGAGGGCAGC 405
Qy 481 CACGTAGTGTGCTGCTGCTCCCGCGCTGAGACACCCATGAGCTCTCACATCCGCTAC 540
Db 406 CACGTGTGTGCTGCTGCTGCTGCCACCTCTCTGGAGCAGCTATGACACCCACATCCGATAT 465
Qy 541 GAGTTGAGCTTCTGGCGGCGCAAGCGCGAGGAGCGTACAGAGGTGGAGATCTCTGGAG 600
Db 466 GAATGAGAGCTGTGCGGAGGCAACCGGCGAGGAGGAGACAAAGGGTGGAGGCTCTGGAA 525
Qy 601 GGCGCGACCGAGTGTGCTGAGCAACTGCGGCGCGGAGCGCTACACTTCGCGCTC 660
Db 526 GGCGCGACTGAGTGTGCTGAGCAACTGCGGCGCGGAGCGCTACACTTCGCTGTT 585
Qy 661 CGCGCGCTATGCTGAGCGCGCTTCCGGCGCTTCTGGAGCGCTGCTCGAGCGCTGTG 720
Db 586 CGAGCGCGATGGCGAGCGAGCTTCAAGCGGATCTCTGAGTGTGCTGAGCGCGCG 645
Qy 721 TCGTGTGCTGAGCGCTAGGAGCTCGGACCCCTCATCTGAGCGCTTCTCCCTCATCTCGTG 780
Db 646 TCACCTACTGAGCGCTAGGAGCTCGGACCCCTCATCTGAGCGCTTCTCTCATCTCGGTC 705
Qy 781 GTCATCTGTGCTGAGCGCTGAGCGCTGCTGCTGCTGCTTCCACCGCGGCTTGAAGCAG 840
Db 706 CTCATCTGCTGTTGCTGAGCGGTTCTGGCCCTGCTGTCCACCGCCCGGACTCTGACGAG 765
Qy 841 AAGATCTGGCTGCGCATCCCGAGCGCCAGAGCGAGTTTGAAGCGCTTTCACCAACCCAC 900
Db 766 AAGATCTGGCTGCGCATCCCAAGCCAGAGCGAGTTTGAAGCGCTTCTTCACCAACCCAC 825
Qy 901 AAGGTAATCTCAGCTGTGCTGTATCAGATGATGCTGCTGCTGCTGAGGAGCGCTCGC 960
Db 826 AAGGTAATCTCAGCTGTGCTGTCTGAGCGCTGATGCTGTCTGTGCTGAGGAGCGCGGC 885
Qy 961 ACCCGCTTTCAGGAGGAGCCACCTGCTTCTCGGAAGTCTCTCTCAGAGCGCTGCTGGGG 1020
Db 886 AGCTCTCTTCTGAGGATCCACCTGCGCCACCTAGAGGTCTCTCTCAGAGCGCCAGCTGGGCA 945
```

```
Qy 1021 ACATGAGCAGCAGTGGAGCGCGGAGCAGATGATGAGGCGCCCTCTGCTGAGCGCAGTGGCG 1080
Db 946 GTGACTCAGGCTGGGACCCAGGCGCAGATGATGAGGCGCCCTTACTGGAGCGGCTGGGC 1005
Qy 1081 AGTGAGCATGCCAGGATACCTATCTGTGCTCTGGAACAATGTTGCTGCCCCCGGAACCCG 1140
Db 1006 AGTGAGCATGCCAGGACACCTACTTGGTATTGGATAAGTGGTGTGCTGCCCCCGGACCCCA 1065
Qy 1141 CCAGTGGAGACCTCCAGGCGCTGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
Db 1066 TGCAGTGAGAACCTCTCAGGCGCTGGGCGAGTGTGGACCTGTGACTATGGATGAAGCT 1125
Qy 1201 TCAGAAAGCATCTCTGCTCATCTGCTTGGCCCTCGAAGCCCGAGCCAGAGGAGCGCTCT 1260
Db 1126 TCAGAAAGCATCTCTGCTCATCTGCTTGGCCCTCGAAGCCCGAGCCAGAGGAGCACCTCA 1185
Qy 1261 GCTGCCAGCTTTGAGTACACTATCTTGGACCCCGAGCTTCCAGCTCTTGCCTCCATGAGCA 1320
Db 1186 CCTTCCAGCTTTGAGTACACTATCTTGGACCCCGAGCTTCCAGCTCTTGCCTCCCGCTCGGCA 1245
Qy 1321 CTGTGCGCTGAGCTGCGCCCTACCCACCCACCTAAAGTACCTTGTGCTGCTGCTGCTGCT 1380
Db 1246 CTGCTCCCGAGCTACTTCCACTTCCACTTCACTTGAAGTACCTATACCTTGTGCTGCT 1305
Qy 1381 GACTCTGGCATCTCAACTGACTACGCTCAGGGAATCCAGGAGCGCCAGGCGGCTTA 1440
Db 1306 GATCTGGCATCTCAACAGATTACAGTTCGGGGGCTCTCAGGAGTCCACGCGGAGCTCA 1365
Qy 1441 TCCGATGGCGCTTCTTCCAGCCCTTATGAGACAGCTTATCCAGCGCTGAGCGCTCTG 1500
Db 1366 TCTGATGGCGCTTCTTCCAGCCCTTATGAGACAGCTTGTGCTCCAGACTCAGAGCTCTG 1425
Qy 1501 CCCCAGCAGCTATGTGCTGCTCTTAG 1527
Db 1426 CATCCCGCTATGTGCGCTGCTCTAG 1452
```

RESULT 8
BX382170/c

LOCUS BX382170 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0D1076YCl3 3-PRIME, mRNA sequence.

ACCESSION BX382170
VERSION BX382170.2 GI:46572612
KEYWORDS EST.

SOURCE
Homo sapiens (human)

Organism: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 1082)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

JOURNAL

On May 8, 2003 this sequence version replaced gi:30449031.

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9443.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?se=CS0D1076AB07NP16c=9443.r.

FEATURES
source

1..1082
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1076YCl3"
/tissue_type="PLACENTA COT 25-NORMALIZED"

Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.jp/) for further details.

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase. and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTAAATTAATCCCGCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
 Host: SOLR.

FEATURES

source

Location/Qualifiers

1..1524
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:2510015H03"
 /db_xref="taxon:10090"
 /clone="2510015H03"
 /tissue_type="liver"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 <1..1377

misc_feature

/note="erythropoietin receptor (MGD|GI:95408, GB|NM_010149, evidence: BLASTN, 99%, match=1523) putative"

ORIGIN

Query Match 62.2%; Score 949.6; DB 3; Length 1524;
 Best Local Similarity 82.0%; Pred. No. 1.4e-205;
 Matches 1131; Conservative 0; Mismatches 244; Indels 5; Gaps 3;
 QY 149 TTCTGTGCTTACCAGCGGTTGGAGGACTTGGTGTCTTGGAGGAGCGGCGAGCT 208
 DB 2 TTCTGTGCTTACCAGCGTGGAGACTTGGTGTCTTGGAGGAGCGGCGAGCT 61
 QY 209 CTGGGTGGGCGCGGCAACTACAGCTTCTCTACAGCTCGAGATGAGCCATGGAGC 268
 DB 62 CCGGATGGAC---TTCAACTACAGCTTCTATACAGCTCGAGGAGTACGAAAGT 118
 QY 269 TGTGTGCTGCTACAGAGCTCCAGCGCTCGTGGTGGTGGTGGTGGTGGTGGTGG 328
 DB 119 CATGTAGCTGCTACAGAGCTCCAGCGTCCAGCGTCCGTCGCTTCTGGTTCAC 178
 QY 329 CTACAGCGGACAGCTCGAGCTTCTGTCGCTTCTAGAGTTTGGGCTCACAGCAGCTCCGCGC 388
 DB 179 CAACAGCGGACATCGAGTTTGTGCGCTCGAGCTCGAGTGCAGGAGCGGTCGGTT 238
 QY 389 CTCGCGGATATACCGTGTCTATCCATCAATGAAGTAGTGTCTCTAGAGCGCGGTTG 448
 DB 239 CTCCTCGCTATACCGCATCATCCATATCAATGAAGTAGTGTCTCTGAGCGCGCGG 298
 QY 449 GGCTGTGTCGCGGCTTGGCTGACGAGCGCGGCTAGTGTGGCTGCTCCCGCG 507
 DB 299 GGCTGTGTCGCGCGCGGCGAGAGAGGCGGACCGAGTGTGGCTGCTCCACCT 358
 QY 508 CTGAGACACCATGACGTCTCACATCCGCTACAGAGTGGAGCTCTCGCGCGGCAAGCGC 567
 DB 359 CCTGGAGCACTATGACCAACCATCCGATATGAGTGGAGCTGTGGAGAGGCAACCGG 418

QY 568 GCAGGAGCGTACAGAGGTTGAGATCTTGAGAGGCGCGCACCGAGTGTGTCTGAGCAAC 627
 DB 419 GCAGGAGGACACAAAGAGGTGGAGTCTCTGGAAGCGGCACCTAGTGTGTCTGAGCAAC 478
 QY 628 CTGGGCGCGGAGCGGCTACCTTCGCGGTCCGCGGCGGTATGGCTGAGCCGAGCTTC 687
 DB 479 CTGGGCGCGGAGC-CGCTACACTTTCGCTGTTCGAGCGGCGCATGGCGGAGCCGAGCTTC 537
 QY 688 GCGGCTTCTGGAGCGCTGCTGGAGCCCTGTGTGCTGTGTGAGCGCTTAGCGACCTGGAC 747
 DB 538 AGCGGATTCGGAGTGCCTGCTGAGCCCGCGTCACTACTGACCGCTAGCGACCTGGAC 597
 QY 748 CCCTCATCTGAGCGCTCTCCCTCATCTCTGCTGTGTGTCTGCTGTGCTGAGCGTCTG 807
 DB 598 CTTCTCATCTGAGCGTGTCTCTCATCTCTGCTGTGTGTGCTGAGCGTGTCTG 657
 QY 808 GCGCTGCTCTCCACCGCGGCTCTGAAGCAGAGATCTGGCGTGCATCCCGAGGCCA 867
 DB 658 GCGCTGCTCTCCACCGCGGCTCTGAGCAGAGATCTGGGCTGCGATCCCAAGGCCA 717
 QY 868 GAGAGCGAGTTTGAAGGCTCTTTCACACCCCAAGGCTAACTTCCAGCTGTGGCTGTAC 927
 DB 718 GAGAGCGAGTTTGAAGGCTCTTTCACACCCCAAGGCTAACTTCCAGCTGTGGCTGTG 777
 QY 928 CAGATGATGGCTGCTGCTGAGGCGCTGCACCCCTTTCAGGAGGAGCCACCTGCT 987
 DB 778 CAGCGTGTGCTGCTGAGGCGCGGCGAGCTCTTTCCTGAGGATTCACCTGCTGCC 837
 QY 988 TCCCTGGAAGTCTCTCAGAGCGCTGCTGGGAGCAGATGAGGAGTGGAGCGCGGACA 1047
 DB 838 CACCTAGAGTCTCTCTCAGAGCAGCTGCGGAGTGACTCAGGCTGGGAGCCAGGGGCA 897
 QY 1048 GATGATGAGGCGCGCTTCTGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1107
 DB 898 GATGATGAGGCGCGCTTCTGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 957
 QY 1108 GTGCTGAGCAAAATGCTGCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1167
 DB 958 GTATTGGATAAGTGTGCTGCGCGGAGCGCGGAGCGCGGAGGAGGAGGAGGAGGAGG 1017
 QY 1168 GGAGTGTGACATAGTGGCATGAGGCTGAGAGGCTCAGAGGATCTCTGCTGCTCATCTGCT 1227
 DB 1018 GGAGTGTGACCGCTGTGACTATGATGAGGCTCAGAGGATCTCTGCTGCTGCTGCTGAC 1077
 QY 1228 TTGCTCTGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1287
 DB 1078 TTGCTCTGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1137
 QY 1288 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1347
 DB 1138 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1197
 QY 1348 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1407
 DB 1198 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1257
 QY 1408 TCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467
 DB 1258 TCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1317
 QY 1468 GAG 1527
 DB 1318 GAG 1377

RESULT 10

CR597800 1320 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0D1021VL06 of Placenta Cot 25-normalized
 DEFINITION of Homo sapiens (human).
 ACCESSION CR597800
 VERSION CR597800.1 GI:50478607
 KEYWORDS HTC; CNSLT_cDNA.

Query Match 56.1%; Score 863.8; DB 5; Length 984;
 Best Local Similarity 98.9%; Pred. No. 4.7e-186;
 Matches 898; Conservative 2; Mismatches 5; Indels 3; Gaps 3;

Qy	1	ATGACACACCTCGGGGGCTCCCTCTGGCCCCAGGTCGGCTCCCTTTGTCTCTCTCGCT	60
Db	79	ATGACACACCTCGGGGGCTCCCTCTGGCCCCAGGTCGGCTCCCTTTGTCTCTCTCGCT	138
Qy	61	GGGGCGCTCGGGCGCCCGCGCTTAACCTCCGGACCCCAAGTTTCGAGAGCAAGCGGCC	120
Db	139	GGGGCGCTCGGGCGCCCGCGCTTAACCTCCGGACCCCAAGTTTCGAGAGCAAGCGGCC	198
Qy	121	TTGCTGCGCGCCCGGGGGCCGGAAGAGCTTCTGTCTTACCGAGCGGTTTCGAGACTTG	180
Db	199	TTGCTGCGCGCCCGGGGGCCGGAAGAGCTTCTGTCTTACCGAGCGGTTTCGAGACTTG	258
Qy	181	GTGTGTTTCGGAGAAAGCGGAGCGCTGGGGTGGGCCCGGGCAACTACAGACTTCTCC	240
Db	259	GTGTGTTTCGGAGAAAGCGGAGCGCTGGGGTGGGCCCGGGCAACTACAGACTTCTCC	318
Qy	241	TACACGCTGAGATCAGCAATGGAAGCTGTCTGCCCTGACCAAGGCTCCACAGGCTCGT	300
Db	319	TACACGCTGAGATCAGCAATGGAAGCTGTCTGCCCTGACCAAGGCTCCACAGGCTCGT	378
Qy	301	GGTGGGTCGCTTCTGTGTTTCGCTGCTACAGCGACGCTCGAGCTTCGTGCCCTTA	360
Db	379	GGTGGGTCGCTTCTGTGTTTCGCTGCTACAGCGACGCTCGAGCTTCGTGCCCTTA	438
Qy	361	GAGTTCGGCTCACAGACGCTCCGGCGTCCGGATATACCGTGTATCCACATCAAT	420
Db	439	GAGTTCGGCTCACAGACGCTCCGGCGTCCGGATATACCGTGTATCCACATCAAT	498
Qy	421	GAAGTAGTGTCTAGACCCCGCTGGGGCTGTGGCGGTTGGCTGACGAGAGCGGC	480
Db	499	GAAGTAGTGTCTAGACCCCGCTGGGGCTGTGGCGGTTGGCTGACGAGAGCGGC	558
Qy	481	CAGGTAGTGTGGCTGCGCGCTGAGACACCCATGAGCTCTCATATCCGCTAC	540
Db	559	CAGGTAGT-TTGGCTGGCTCCCGCGCTGAGACACCCATGAGCTCTCATATCCGCTAC	617
Qy	541	GAGGTGAGCTTCGGCGGCAACGGCGCA-GGGAGCGTACAGAGGTTGGAGATCCTTGA	599
Db	618	GAGGTGAGCTTCGGCGGCAACGGCGGAGGGAGCGTACAGAGGTTGGAGATCCTTGA	677
Qy	600	GGGCGCCACCGAGTGTGCTGAGCAACTCGGGGGCGGAGCGGCTACACTTCGCGCT	659
Db	678	GGGCGCCACCGAGTGTGCTGAGCAACTCGGGGGCGGAGCGGCTACACTTCGCGCT	737
Qy	660	CGCGCGCGTATGCTGAGCGAGCTTCGGGGCTTTTGGAGCGCGCTGCTGGAGCGCTGT	719
Db	738	CGCGCGCGTATGCTGAGCGAGCTTCGGGGCTTTTGGAGCGCGCTGCTGGAGCGCTGT	797
Qy	720	GTGCTGTGACGCTTAGCGACCTTGACCCCTCATCTGACGCTCTCCCTCATCTCTGT	779
Db	798	GTGCTGTGACGCTTAGCGACCTTGACCCCTCATCTGACGCTCTCCCTCATCTCTGT	857
Qy	780	GGTATCTGCTGCTGCTGACCGTCTCGCGCTCTTCCACCGCGGGCTCTGAAGCA	839
Db	858	GGTATCTGCTGCTGCTGACCGTCTCGCGCTCTTCCACCGCGGGCTCTGAAGCA	917
Qy	840	GAAGATCTGGCTGACATCCGAGCCAGAGCGGAGTTGAAGGCTCTTCCACCA	899
Db	918	GAAGATCTGGCTTGGATCGAGGCCAGAGCGGAGTTGAAGGCTCTTCCACCA	976
Qy	900	CAAGGGTA 907	
Db	977	CAAGGGTA 984	

RESULT 12
 AL553257
 LOCUS
 DEFINITION Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 1079 bp mRNA linear EST 30-MAR-2004

RESULT 12	AL553257	LOCUS	DEFINITION
-----------	----------	-------	------------

Db 463 GAAGTAGTCTCCAGAGCGCCCGTGGGGCTGGTGGCGGTTGGCTGACGAGCGG 522

Qy 480 CCACGTAGTGTGGCTGGCTCCCGCCCTGAGACACACCATGAGCTCTCACATCCGCTA 539

Db 523 CCACGTAGTGTGGCTGGCTCCCGCCCTGAGACACACCATGAGCTCTCACATCCGCA 582

Qy 540 CGAGGTGACGTCTCGGCGGCAAGCGGCGAGGAGCGTACAGAGGTGGAGATCCTGGGA 599

Db 583 CGAGGTGACGTCTCGGCGGCAAGCGGCGAGGAGCGTACAGAGGTGGAGATCCTGGGA 642

Qy 600 GGGCGCGACCGAGTGTGTGAGCAACTCCGCGGCGGAGCGCTACACCTTCGCGCGT 659

Db 643 GGGCGCGACCGAGTGTGTGAGCAACTCCGCGGCGGAGCGCTACACCTTCGCGCGT 702

Qy 660 CCGCGCGGTATGGCTGAGCGGAGCTTCGCGCGCTTCCTGGAGCGCTGGTGGAGCGCTGT 719

Db 703 CCGCGCGGTATGGCTGAGCGGAGCTTCGCGCGCTTCCTGGAGCGCTGGTGGAGCGCTGT 762

Qy 720 GTCCCTGTGAGCGGTAGGACCTGGACCCCTCATCTGAGCGTCTCCCTCATCTCGT 779

Db 763 GTCCCTGTGAGCGGTAGGACCTGGACCCCTCATCTGAGCGTCTCCCTCATCTCGT 822

Qy 780 GGTCTATCTGTGTGTGAGCGGCTTCGCGCTCTCTCCACCGCGGCTCTGAAGCA 839

Db 823 GGTCTATCTGTGTGTGAGCGGCTTCGCGCTCTCTCCACCGCGGCTCTGAAGCA 882

Qy 840 GAAGATCTGGCTGGCTCCCGAGCCAGAGAGCGAGTTTGAAGCGCTCTTACACCCCA 899

Db 883 GAAGATCTGGCTGGCTCCCGAGCCAGAGAGCGAGTTTGAAGCGCTCTTACACCCCA 940

Qy 900 CAAGGGTAACTTCAGCTGTGGCTGTACCAAGATGATGGCTG 942

Db 941 CAAGGGTAACTTCAGCTGTGGCTGTACCAAGATGATGGCTG 981

RESULT 13

BX335579 1005 bp mRNA linear EST 08-APR-2004
 LOCUS BX335579 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1016YL05 5-PRIME, mRNA sequence.

ACCESSION BX335579

VERSION BX335579.2 GI:46281994

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1005) Li, W.B., Gruber, C., Jesse, J., and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 2, 2003 this sequence version replaced gi:30345434.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

9443.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0D1016CF03QP1&c=9443.r.

FEATURES

source

1..1005

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1016YL05"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 53.2%; Score 812.2; DB 5; Length 1005;
 Best Local Similarity 96.1%; Pred. No. 2.7e-174;
 Matches 864; Conservative 18; Mismatches 12; Indels 5; Gaps 5;

Qy 1 ATGACACACCTCGGGCGTCCCTCTGCCCCAGGTGGCTCCCTTTGCTCTCCTCGCTCGCT 60

Db 93 ATGACACACCTCGGGCGTCCCTCTGCCCCAGGTGGCTCCCTTTGCTCTCCTCGCTCGCT 152

Qy 61 GGGCGCGCTGGCGCGCCCGCCCTAACTCCCGGACCCCAAGTTTCGAGAGCAAGCGGCC 120

Db 153 GGGCGCGCTGGCGCGCCCGCCCTAACTCCCGGACCCCAAGTTTCGAGAGCAAGCGGCC 212

Qy 121 TTGCTGGCGCCCGGGGGCCGGAAGACTTCTGTGTTTACCCGAGCGGTTGGAGGACTTG 180

Db 213 TTGCTGGCGCCCGGGGGCCGGAAGACTTCTGTGTTTACCCGAGCGGTTGGAGGACTTG 272

Qy 181 GTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGGCCCGGCAACTACAGTCTCTCC 240

Db 273 GTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGGCCCGGCAACTACAGTCTCTCC 332

Qy 241 TACCAGCTCGAGGATGAGCCATGAAAGCTGTGTGCTTGCACCAAGCTCCACCGGCTCGT 300

Db 333 TACCAGCTCGAGGATGAGCCATGAAAGCTGTGTGCTTGCACCAAGCTCCACCGGCTCGT 392

Qy 301 GGTGCGGTGGCTTCTGGTTCGCTGCTTACAGCCGACACGTCGAGCTTCGTCGCCCTA 360

Db 393 GGTGCGGTGGCTTCTGGTTCGCTGCTTACAGCCGACACGTCGAGCTTCGTCGCCCTA 452

Qy 361 GAGTTCGCGCTCACAGCAGCTCCGCGGCTCCCGGATATCACCGTGTATCCACATCAAT 420

Db 453 GAGTTCGCGCTCACAGCAGCTCCGCGGCTCCCGGATATCACCGTGTATCCACATCAAT 512

Qy 421 GAAGTAGTCTCTAGACGCCCGCTGGGGCTGGTGGCGGTTGGCTGTACGAGAGCGGC 480

Db 513 GAAGTAGTCTCTAGACGCCCGCTGGGGCTGGTGGCGGTTGGCTGTACGAGAGCGGC 572

Qy 481 CAGTAGTCTGGCTGGCTCCCGCGCTGAGACACCATGACGCTTCACATCCGCTAC 540

Db 573 CAGTAGTCTGGCTGGCTCCCGCGCTGAGACACCATGACGCTTCACATCCGCTAC 632

Qy 541 GAGGTGAGCTCTCGCGCGCAACCGCGAGCGGTACAGAGGCTGAGATCTCTGGAG 600

Db 633 GAGGTGAGCTCTCGCGCGCAACCGCGAGCGGTACAGAGGCTGAGATCTCTGGAG 691

Qy 601 GCGCGCACCGAGTGTGTGTGAGCAACCTGCGGGCGCGGACCGCTTACACCTTCGCGCTC 660

Db 692 GCGCGCACCGAGTGTGTGTGAGCAACCTGCGGGCGCGGACCGCTTACACCTTCGCGCTC 751

Qy 661 CCGCGCGGTATGGCTGAGCGGAGCTTCGGGGCTTCGGAGCGGCTTCGGAGCGGCTGTG 720

Db 752 CCGCGCGGTATGGCTGAGCGGAGCTTCGGGGCTTCGGAGCGGCTTCGGAGCGGCTGTG 811

Qy 721 TCGCTCTGACGCTAGCGACCTGAGACCCCTCATCTGACGCTCTCCCTCATCTCTCGT 780

Db 812 TCGCTCTGACGCTAGCGACCTGGA-CCCTCATCTGACGCTCTCCCTCATCTCTCGT 870

Qy 781 GTCATCTCGTGTGCTGACCGCTGCTGCTCTCCACCGCGCGGCTCTTGAAGCAG 840

Db 871 GTCATCTCGTGTGCTGACCGCTGCTGCTCTCCACCGCGCGGCTCTTGAAGCAG 929

Qy 841 AGATCTGCGCTGGATCCGAGCGGAGGAGGAGTTTGAAGGCTCTTTCACACCCA 899

Db 930 -AGRTTTGSGCTKGSATYCCGAG-CCAGARARCGAGTTTAARGCYYYTACCACCMCAA 986

RESULT 14

BX448343/c

LOCUS

1051 bp mRNA linear EST 05-MAY-2004

BX448343

```

DEFINITION BX448343 Homo sapiens FETAL LIVER Homo sapiens cDNA clone
CSODM009YC22 3-PRIME, mRNA sequence.
ACCESSION BX448343
VERSION BX448343.2 GI:47052187
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1051)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31019923.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 9443.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSOAM009BBI1NP1&c=9443.r.
FEATURES
Location/Qualifiers
1..1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODM009YC22"
/tissue_type="FETAL LIVER"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL LIVER"
/notes="Organ: liver; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 51.9%; Score 792.2; DB 5; Length 1051;
Best Local Similarity 97.7%; Pred. No. 9.9e-170;
Matches 820; Conservative 4; Mismatches 12; Indels 3; Gaps 2;
Qy 689 CGCGCTTCTGAGCGCTGTCGAGCGCTGTCGCTGCTCAGCCCTAGCGACCTGGACC 748
Db 1042 CGCGCTTTTGAGCGCTGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
Qy 749 CCCTCATCTGACGCTCTCCCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808
Db 985 CCCTCATCTGACGCTCTCCCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
Qy 809 CGCTGCTCTCCACCGCGCGCTCTGAAGCAGAGAGATGCGCTTGGCATCCCGAGCCAG 869
Db 925 CGCTGCTCTCCACCGCGCGCTCTGAAGCAGAGAGATGCGCTTGGCATCCCGAGCCAG 866
Qy 869 AGACGAGTTTGAAGGCTCTTCCACCCACAGAGGTTAACTCCAGCTGTGCTGTACC 928
Db 865 AGACGAGTTTGAAGGCTCTTCCACCCACAGAGGTTAACTCCAGCTGTGCTGTACC 806
Qy 929 AGAATGATGGCTGCTGTGTTGGAGCCCTCTCACCCCTTTCACGAGAGACCCACCTGCTT 988
Db 805 AGAATGATGGCTGCTGTGTTGGAGCCCTCTCACCCCTTTCACGAGAGACCCACCTGCTT 746
Qy 989 CCCTGGAAGTCTCTCAGAGCGCTGCTGTTGGGGAAGATGACAGCAGTGGAGCCGGGACAG 1048
Db 745 CCCTGGAAGTCTCTCAGAGCGCTGCTGTTGGGGAAGATGACAGCAGTGGAGCCGGGACAG 686
Qy 1049 ATGATAGGGGCCCCCTGCTGAGCGCAGTGGGCGAGTGCAGTACCCAGATACCTATCTGG 1108
Db 685 ATGATAGGGGCCCCCTGCTGAGCGCAGTGGGCGAGTGCAGTACCCAGATACCTATCTGG 626

```

```

Qy 1109 TGCTGCACAAATGTTGCTGCTGCCCCGGAACCCGCCCCAGTGAAGGACCTCCAGGGCCTGGT 1168
Db 625 TGCTGCACAAATGTTGCTGCTGCCCCGGAACCCGCCCCAGTGAAGGACCTCCAGGGCCTGGT 566
Qy 1169 GCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAGAAGCATCTCTGCTCATCTGCTT 1228
Db 565 GCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAGAAGCATCTCTGCTCATCTGCTT 506
Qy 1229 TGGCCTCGAAGCCAGCCAGAGGAGGAGCTCTGCTGCCAGCTTTGAGTACACATATCTCTGG 1288
Db 505 TGGCCTCGAAGCCAGCCAGAGGAGGAGCTCTGCTGCCAGCTTTGAGTACACATATCTCTGG 446
Qy 1289 ACCCCAGCTCCAGCTCTTGGCTGCCATGGACACTGTGCCCTGAGTGCCTCCACCCCTTATG 1348
Db 445 ACCCCAGCTCCAGCTCTTGGCTGCCATGGACACTGTGCCCTGAGTGCCTCCACCCCTTATG 386
Qy 1349 CCCACCTTAAAGTACCTGTACCTTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1408
Db 385 CCCACCTTAAAGTACCTGTACCTTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 326
Qy 1409 CAGGGGACTCCAGAGGAGCCCAAGGGGGCTTATCCGATGGCCCTACTCCCAACCCCTTATG 1468
Db 325 CAGGGGACTCCAGAGGAGCCCAAGGGGGCTTATCCGATGGCCCTACTCCCAACCCCTTATG 266
Qy 1469 AGAACAGCTTATCCAGCGCTGAGCCCTGTGCCCCCAGCTATGTGGCTGCTCTTATG 1527
Db 265 AGAACAGCTTATCCAGCGCTGAGCCCTGTGCCCCCAGCTATGTGGCTGCTCTTATG 207
RESULT 15
AY414847 1115 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes EPOR gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY414847
VERSION AY414847.1 GI:39770806
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 1115)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL 14671302
REFERENCE 2 (bases 1 to 1115)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..1115
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1115
/gene="EPOR"
/locus_tag="HCM5360"
ORIGIN
Query Match 51.7%; Score 789; DB 9; Length 1115;
Best Local Similarity 71.0%; Pred. No. 5.4e-169;

```

```
Matches 792; Conservative 0; Mismatches 323; Indels 0; Gaps 0;
Qy 116 CCGCCTTGCTGGCGGCCCGGGGCCGGAAGAGCTTCTGTGCTTTCACCGAGCGGTTGGAGG 175
Db 1 CCGCCTTGCTGGCGGCCCGGGGCCGGAAGAGCTTCTGTGCTTTCACCGAGAGGTTGGAGG 60
Qy 176 ACTTGTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGCCCGGGCAACTACAGCT 235
Db 61 ACTTGTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGCCCGGGCAACTACAGCT 120
Qy 236 TCTCTTACCAGCTCGAGGATGAGCCATGGAAGTGTGTCGCTGCACACAGGCTCCCAAG 295
Db 121 TCTCTTACCAGCTCGAGGATGAGCCATGGAAGTGTGTCGCTGCACACAGGCTCCCAAG 180
Qy 296 CTCTGTGTCGCTGCGCTTCTGTGTTCTGCTACAGCCGACACGTGAGCTTCGTGC 355
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
Qy 356 CCTAGAGTTGGCGTCA CAGAGCTTCGGCGCTCCGGCGATATCACCGTGTATCCACA 415
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
Qy 416 TCAATGAAGTAGTCTCTAGACGCCCGCTGGGCTGGTGGCGGTTGGCTGACGAGA 475
Db 301 TCNNNNNNNNNNCTCTTAGACGCCCGCTGGGCTGGTGGCGGTTGGCTGACGAGA 360
Qy 476 GCGGCCACGTAGTGTGCTGCTCCCGCCCTGAGACACCCATGACGTCTCACATCC 535
Db 361 GCGGCCACGTAGTGTGCTGCTCCCGCCCTGAGACACCCATGACGTCTCACATCC 420
Qy 536 GCTACGAGGTGACGTCTCGGCGGGAACGGCGCAGGAGCGTACAGAGGTGGAGATCC 595
Db 421 GCTACNAGGTGACGNNCGCGCAACGGCGCAGGAGCGTACAGAGGNNNNNNNNNN 480
Qy 596 TGGAGGGCCGACCGAGTGTGTCTGACGAACTCGGGGGCCGAGCGCTACACTTCG 655
Db 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540
Qy 656 CCGTCGCGCGCGTATGGCTGAGCGCGCTTCGGCGGCTTCTGGAGCGCTGCTCGGAGC 715
Db 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
Qy 716 CTGTGCTGCTGAGCGCTAGCGACTGGAGCCCTCATCTCTGAGCGCTCTCCCTCATCC 775
Db 601 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660
Qy 776 TCGTGTCTATCTGTGTGTCTGACCGCTGCTCGCGCTCTCCACCGCGGGCTCTGA 835
Db 661 TCGTGTCTATCTGTGTGTCTGACCGCTGCTCGCGCTCTCTCCACCGCGNNNNNGA 720
Qy 836 AGCAGAAGATCTGGCTGGCATCCCGAGCCACAGAGCGAGTTTGAAGGCTCTTTCACCA 895
Db 721 AGCAGAAGANNNGCTTGGCATCCCGAGCCACAGAGCGAGTTTGAAGGCTCTTTCACCA 780
Qy 896 CCCACAAGGGTAACCTTCAGCTGTGGCTGTACCAAGATGATGGCTGCTGTGGTGGAGCC 955
Db 781 CCCACAAGGGTAACCTTCAGCTGTGGCTGTACCAAGATGATGGCTGCTGTGGTGGAGCC 840
Qy 956 CTTGCAACCCCTTTCAGCGAGGACCCACTGCTTCTCCCTGGAAGTCTCTCAGAGCGCTGT 1015
Db 841 CTTGCAACCCCTTTCAGCGAGGACCCACTGCTTCTCCCTGGAAGTCTCTCAGAGCGCTGT 900
Qy 1016 GGGGACGATGAGGAGTGGAGCGGGGACAGATGATGAGGGCCCTCTGTGGAGCCAG 1075
Db 901 GGGGACGATGAGGAGTGGAGCGGGGACAGATGATGAGGGCCCTCTGTGGAGCCAG 960
Qy 1076 TGGGAGTGAAGATGCCAGGATACCTATCTGGTGTCTGGACAAATGGTGTGCCCCGGA 1135
Db 961 TGGGAGTGAAGATGCCAGGATACCTATCTGGTGTCTGGACAAATGGTGTGCCCCGGA 1020
Qy 1136 ACCCGCCAGTCAGGACCTCCAGGGCTGTGGCAGTGTGGACATAGTGGCCATGGATG 1195
Db 1021 ACCCGCCAGTCAGGACCTCCAGGGCTGTGGCAGTGTGGACATAGTGGCCATGGATG 1080
```

Qy 1196 AAGGCTCAGAAAGCATCTCTGCTCATCTGCTTTG 1230
Db 1081 AAGGCTCAGAAAGCATCTCTGCTCATCTGCTTTG 1115

Search completed: April 21, 2005, 04:08:44
Job time : 5271 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 19:27:26 ; Search time 849.5 Seconds
(without alignments)
10640.901 Million cell updates/sec

Title: US-09-016-159D-4
Perfect score: 1527
Sequence: 1 atggaccactcgggcgctc.....gctatgtgcttgccttag 1527

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1527	100.0	1527	2	AAQ82990 Human ery
2	1527	100.0	1848	13	Adq83546 Human tum
3	1527	100.0	1849	12	Ado05723 Human ery
4	1527	100.0	1849	12	Adp10351 Reference
5	1527	100.0	1849	13	Acn40465 Tumour-as
6	1527	100.0	1883	2	AAQ05748 EPO recep
7	1525.4	99.9	1527	6	AD151519 Human ery
8	1525.4	99.9	1818	2	AAQ81892 Human ery
9	1520.6	99.6	1585	10	AdE28676 Human NOV
10	1520.6	99.6	1585	12	Adm93421 Human NOV
11	1515.8	99.3	1818	2	AAQ53995 Human EPO
12	1312.2	85.9	1317	3	Aaz49634 Truncated
13	1307.4	85.6	1317	3	Aaz49636 Mutant R1
14	1209	79.2	1733	10	AdE28670 Human NOV
15	1209	79.2	1733	10	AdE28672 Human NOV
16	1209	79.2	1733	12	Adm93417 Human NOV
17	1209	79.2	1733	12	Adm93415 Human NOV
18	1204.2	78.9	1435	10	AdE28674 Human NOV
19	1204.2	78.9	1435	12	Adm93419 Human NOV
20	1080.6	70.8	1740	2	AAQ05747 EPO recep

21	1075.8	70.5	1741	2	AAQ53994	AAQ53994 MEL EPO r
22	1075.8	70.5	1741	2	AAQ44854	AAQ44854 Mouse sol
23	1075.8	70.5	1741	2	AAQ81891	AAQ81891 Mouse ery
24	1024	67.1	1024	10	ABZ83469	ABZ83469 Toxicolog
25	811.2	53.1	5068	13	ADR70427	ADR70427 Vector pa
26	811.2	53.1	5068	13	ADR70421	ADR70421 Vector pa
27	810.4	53.1	5683	6	AAAL41116	AAAL41116 Plasmid p
28	810.4	53.1	6122	6	ABN84078	ABN84078 Plasmid p
29	810.4	53.1	6122	6	AAAL41112	AAAL41112 Plasmid p
30	810.4	53.1	6122	6	AAAL41115	AAAL41115 Plasmid p
31	810	53.0	1121	8	ABZ23425	ABZ23425 Nucleotid
32	810	53.0	4896	13	ADR70403	ADR70403 Vector pa
33	809.6	53.0	1056	8	ABZ23429	ABZ23429 Nucleotid
34	809.6	53.0	5059	13	ADR70409	ADR70409 Vector pa
35	809.2	53.0	1131	8	ABZ23424	ABZ23424 Nucleotid
36	808.8	53.0	6125	6	AAAL41114	AAAL41114 Plasmid p
37	808	52.9	1059	8	ABZ23428	ABZ23428 Nucleotid
38	808	52.9	5059	13	ADR70415	ADR70415 Vector pa
39	750.4	49.1	4990	2	AAT48800	AAT48800 Plasmid m
40	750	49.1	750	2	AAV04434	AAV04434 Erythropo
41	750	49.1	750	2	AAZ30854	AAZ30854 Human ery
42	748.2	49.0	6256	6	AAAL43173	AAAL43173 pCAGGS-hs
43	741	48.5	5565	6	AAAL43171	AAAL43171 Human exp
44	700.4	45.9	1666	2	AAx58148	AAx58148 CadC-fusi
45	602.8	39.5	4883	6	ABQ73377	ABQ73377 Vector pr

ALIGNMENTS

RESULT 1		AAQ82990		AAQ82990 standard; cDNA; 1527.BP.	
ID	AAQ82990	standard; cDNA; 1527.BP.			
XX	AC	AAQ82990;			
XX	AC	AAQ82990;			
DT	25-MAR-2003	(revised)			
DT	07-OCT-1995	(first entry)			
XX	XX	Human erythropoietin receptor cDNA.			
XX	XX	Erythropoietin receptor; extracellular domain; plasmid pJYL26; ds.			
KW	XX	Homo sapiens.			
OS	XX	Homo sapiens.			
Key	Key	Location/Qualifiers			
FT	CDS	1..1527			
FT	FT	/*tag= a			
FT	misc_feature	25..250			
FT	FT	/*tag= c			
FT	misc_feature	73..750			
FT	FT	/*tag= d			
FT	misc_feature	73..87			
FT	FT	/*tag= b			
FT	FT	/*tag= "extracellular domain"			
FT	FT	/*tag= "extracellular domain"			
FT	FT	/*tag= "forward primer AAQ82991 site"			
XX	XX	WO9505469-A1.			
XX	XX	23-FEB-1995.			
PD	PD	15-AUG-1994;			
XX	XX	94WO-US009298.			
XX	XX	16-AUG-1993;			
XX	XX	93US-00106815.			
XX	XX	(LEEJ/) LEE J Y.			
XX	XX	Lee JY;			
XX	XX	WPI; 1995-098767/13.			
XX	XX	P-PSDB; AAR70032.			

PT New pure human erythropoietin receptor fragment - obtd. by expression as
PT a fusion protein having a thrombin proteolytic cleavage site.
XX
PS Disclosure; Page 27-29; 42pp; English.

The cDNA encodes human full-length erythropoietin receptor (EPO-R). Primers AQ82991 and AQ82992 are used to isolate fragments (nt 73-750 and 25-250) representing the extracellular domain of the receptor. The fragments are cloned in vector plasmid pGEX-2T, resulting in vector pY126, which encodes the extracellular domain as a fusion protein with glutathione-S-transferase. The extracellular domain of EPO-R is used for investigating the structure of the EPO-R and for identifying factors involved in regulating differentiation and proliferation mechanisms in erythroid progenitor cells. It can also be used for identifying and quantitating EPO and EPO-R as well as in understanding haematopoietic malignancy and some cardiovascular system disorders. (Updated on 25-MAR-2003 to correct PN field.)

Seq	Sequence	1527 BP;	246 A;	514 C;	470 G;	297 T;	0 U;	0 Other;
	Query Match	100.0%;	Score 1527;	DB 2;	Length 1527;			
	Best Local Similarity	100.0%;	Pred. No. 0;					
	Matches 1527;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
Qy	1	ATGACACACCTCGGGGCGTCCTCTTGCGGCCACAGTTCGGCTTCCTTTGTCTCTCTGCTCGCT	60					
Db	1	ATGACACACCTCGGGGCGTCCTCTTGCGGCCACAGTTCGGCTTCCTTTGTCTCTCTGCTCGCT	60					
Qy	61	GGGGCGGCTGGGCGGCCCGCCCTAACTCCCGGACCCCAAGTTCGAGAGCAAAAGCGGCC	120					
Db	61	GGGGCGGCTGGGCGGCCCGCCCGCTAACTCCCGGACCCCAAGTTCGAGAGCAAAAGCGGCC	120					
Qy	121	TTGTGGCGGCCCGGGGCGCCGAAAGAGCTTCTGTCTTTCACGACGGTTGGAGACTTG	180					
Db	121	TTGTGGCGGCCCGGGGCGCCGAAAGAGCTTCTGTCTTCAACGAGCGTTGGAGACTTG	180					
Qy	181	GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCGCGGCAACTACAGCTTCCTCC	240					
Db	181	GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCGCGGCAACTACAGCTTCCTCC	240					
Qy	241	TACCAGCTTCAGAGATGAGCCATGGAAGCTGTCTCGCTTCGACACAGGCTCCACGGCTCGT	300					
Db	241	TACCAGCTTCAGAGATGAGCCATGGAAGCTGTCTCGCTTCGACACAGGCTCCACGGCTCGT	300					
Qy	301	GGTGGGTGCGCTTCTGTGTTCGTGCTACAGCCGACAGTTCGAGCTTCGTGGCCCCCTA	360					
Db	301	GGTGGGTGCGCTTCTGTGTTCGTGCTACAGCCGACAGTTCGAGCTTCGTGGCCCCCTA	360					
Qy	361	GAGTTGGCGGTCACAGCAGCTCCGGCGCTCCGGGATATACCGGTCTACCATCAAT	420					
Db	361	GAGTTGGCGGTCACAGCAGCTTCGGCGCTTCGGGATATACCGGTCTATCCACATCAAT	420					
Qy	421	GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGGTGGCGCGGTTGGCTGACGAGCGGC	480					
Db	421	GAAGTAGTGCTCCTAGACGCCCCCGTGGGGCTGGTGGCGCGGTTGGCTGACGAGCGGC	480					
Qy	481	CACGTAGTGTGGCTGCTCGCGCGCTCAGACACCCATGACGTCTCACATCCGCTAC	540					
Db	481	CACGTAGTGTGGCTGCTCGCGCGCTCAGACACCCATGACGTCTCACATCCGCTAC	540					
Qy	541	GAGGTGACAGTCTCGGCGGCAACGGCGCAGGAGCGTACAGAGGTTGGAGATCCTGGAG	600					
Db	541	GAGGTGACAGTCTCGGCGGCAACGGCGCAGGAGCGTACAGAGGTTGGAGATCCTGGAG	600					
Qy	601	GGCGCACCGAGTGTGTCTGAGCAACTCGGGGCGGACCGCTACACCTTCGCCGTC	660					
Db	601	GGCGCACCGAGTGTGTCTGAGCAACTCGGGGCGGACCGCTACACCTTCGCCGTC	660					
Qy	661	CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCTCTGGAGGCTGTTCGGAGCGCTGTG	720					
Db	661	CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCTCTGGAGGCTGTTCGGAGCGCTGTG	720					
Qy	721	TCGCTGTGTAGCGCTAGCGACTCGGACCCCTCATCTCTGAGCGCTCTCCCTCATCTCGTG	780					

22-JUL-2004.
 15-OCT-2003; 2003WO-US029126.
 18-OCT-2002; 2002US-0418988P.
 (GETH) GENENTECH INC.
 (WUTD) WU T D.
 (ZHOU) ZHOU Y.
 Wu TD, Zhou Y;
 WPT; 2004-534300/51.
 New nucleic acid molecule and encoded polypeptide, for diagnosing,
 preventing or treating cell proliferative disorders such as cancer.
 Claim 1; SEQ ID NO 360; 5504pp; English.
 The present invention describes an isolated tumour-associated antigenic
 target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 (c). Also described: (1) an expression vector comprising the above
 nucleic acid; (2) a host cell comprising the above expression vector; (3)
 a process for producing a polypeptide; (4) an isolated polypeptide
 comprising: (a) an amino acid sequence encoded by any of the above
 nucleotide sequences; (b) an amino acid sequence encoded by the full-
 length coding region of the above nucleotide sequences; or (c) a sequence
 having at least 80% identity to (a) or (b); (5) a chimeric polypeptide
 comprising the above polypeptide fused to a heterologous polypeptide; (6)
 an isolated antibody that binds to the above polypeptide; (7) a process
 for producing the antibody; (8) an isolated oligopeptide that binds to
 the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 binding organic molecule that binds to the above polypeptide; (10) a
 composition of matter comprising the above (chimeric) polypeptide,
 antibody, oligopeptide or TAT binding organic molecule, in combination
 with a carrier; (11) an article of manufacture comprising a container and
 the composition of matter contained within the container; (12) methods of
 inhibiting the growth of a cell that expresses the above protein, where
 the growth of the cell is at least in part dependent upon a growth
 potentiating effect of the above protein; (13) a method of
 therapeutically treating a mammal having a cancerous tumour comprising
 cells that express the above protein; (14) a method of determining the
 presence of a protein in a sample suspected of containing the protein
 described above; (15) methods of diagnosing the presence of a tumour in a
 mammal; (16) a method for treating or preventing a cell proliferative
 disorder associated with increased expression or activity of the above
 protein; and (17) a method of binding an antibody, oligopeptide or
 organic molecule to a cell that expresses the protein described above.
 The TAT sequences have cytostatic activities, and can be used in gene
 therapy. The composition and methods are useful for diagnosing,
 preventing or treating cancer. The composition is also used for preparing
 a medicament for the therapeutic treatment or diagnostic detection of a
 cell proliferative disorder or cancer. The present sequence represents a
 human TAT cDNA sequence from the present invention.

257 TTGCTGGCGCCCGGGGCGCCGGAAGAGCTTCTGTGCTTACCGAGCGGTTGGAGGACTTG 316
 181 GTGTGTTTCTGGGAGGAAGCGGAGCGCTGGGTGGGCGCGGCAACTACAGCTTCTCC 240
 317 GTGTGTTTCTGGGAGGAAGCGGAGCGCTGGGTGGGCGCGGCAACTACAGCTTCTCC 376
 241 TACCAAGCTCAGGATGAGCCATGGAAAGCTGTGTGCTTGCACACAGGCTCCACCGGCTCGT 300
 377 TACCAAGCTCAGGATGAGCCATGGAAAGCTGTGTGCTTGCACACAGGCTCCACCGGCTCGT 436
 301 GGTGGGCTGGCTTCTGGTGGCTTACAGCGGACAGCGTCCGTCGCTTCCGTCCTA 360
 437 GGTGGGCTGGCTTCTGGTGGCTTACAGCGGACAGCGTCCGTCGCTTCCGTCCTA 496
 361 GAGTTGCGGCTCACAGCAGCTCCGCGGCTCCGCGGATATCACCGTGTCTCATCATCAAT 420
 497 GAGTTGCGGCTCACAGCAGCTCCGCGGCTCCGCGGATATCACCGTGTCTCATCATCAAT 556
 421 GAAGTAGTCTCTTACAGCAGCTCCGCGGCTGGTGGCGGCTTGGCTGACAGAGCGGC 480
 557 GAAGTAGTCTCTTACAGCAGCTCCGCGGCTGGTGGCGGCTTGGCTGACAGAGCGGC 616
 481 CAGCTAGTGTGGGCTGGCTCCGCGGCTGAGACACCATGACGCTCTCATCTCCGCTAC 540
 617 CAGCTAGTGTGGGCTGGCTCCGCGGCTGAGACACCATGACGCTCTCATCTCCGCTAC 676
 541 GAGTTGAGCTCTCGCGCGGCAACCGCGGAGGCTACAGAGGCTGAGATCTCTGGAG 600
 677 GAGTTGAGCTCTCGCGCGGCAACCGCGGAGGCTACAGAGGCTGAGATCTCTGGAG 736
 601 GCGCGCACCGAGTGTGTGCTGAGCAACCTGCGGGCGGAGCGGCTACACCTTCCGCGTC 660
 737 GCGCGCACCGAGTGTGTGCTGAGCAACCTGCGGGCGGAGCGGCTACACCTTCCGCGTC 796
 661 CCGCGCGCTATGGCTGAGCGGAGCTTCCGCGGCTTCTGAGCGGCTGCTGGAGCGCTGTG 720
 797 CCGCGCGCTATGGCTGAGCGGAGCTTCCGCGGCTTCTGAGCGGCTGCTGGAGCGCTGTG 856
 721 TCGCTGCTACCGCTAGCGACCTGAGACCCCTCATCTGACGCTCTCCCTCATCTCTCGTG 780
 857 TCGCTGCTACCGCTAGCGACCTGAGACCCCTCATCTGACGCTCTCCCTCATCTCTCGTG 916
 781 GTCTCTCTGGTGTGCTGACGCTGCTGCGCTGTCTTCCACCGCGGCGCTCTGAAGCAG 840
 917 GTCTCTCTGGTGTGCTGACCGCTGCTGCGCTGTCTTCCACCGCGGCGCTCTGAAGCAG 976
 841 AAGATCTGGCTGGCATCCGAGCGGAGCGAGGCTTGAAGGCTCTTACACCCAC 900
 977 AAGATCTGGCTGGCATCCGAGCGGAGCGAGGCTTGAAGGCTCTTACACCCAC 1036
 901 AAGGCTAAGCTTCCAGCTGTGGCTGTACCAAGATGATGGCTGTGTGGAGCGGCTTGC 960
 1037 AAGGCTAAGCTTCCAGCTGTGGCTGTACCAAGATGATGGCTGTGTGGAGCGGCTTGC 1096
 961 ACCCGCTTCCAGGAGACCCAGCTGCTTCCCTGGAGAGTCTCTCAGAGCGCTGTGGGG 1020
 1097 ACCCGCTTCCAGGAGACCCAGCTGCTTCCCTGGAGAGTCTCTCAGAGCGCTGTGGGG 1156
 1021 ACGATCGAGGAGTGGAGCGGCGGAGCAGATGATGAGGCGGCTTCTGAGCGGCTTGC 1080
 1157 ACGATCGAGGAGTGGAGCGGCGGAGCAGATGATGAGGCGGCTTCTGAGCGGCTTGC 1216
 1081 AGTGAGCATGCCAGGATACCTATCTGTGCTGGCAAAATGGTGTGTGCGCGGAAACCG 1140
 1217 AGTGAGCATGCCAGGATACCTATCTGTGCTGGCAAAATGGTGTGTGCGCGGAAACCG 1276
 1141 CCCAGTAGGAGCTCCAGGCGCTGGTGGCAGTGTGGAATAGTGCGGCTATGATGAAGGC 1200
 1277 CCCAGTAGGAGCTCCAGGCGCTGGTGGCAGTGTGGAATAGTGCGGCTATGATGAAGGC 1336
 1201 TCAGAGCATCTCTCTGCTCATCTGCTTGGCTTGAAGCGGCTTGAAGCGGCTTCT 1260

Query Match 100.0%; Score 1527; DB 13; Length 1848;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGACACACTCGGGGCTCCCTTGGCCCCAGGTGGGCTCCCTTGTCTCTCTGCTCGCT 60
 137 ATGACACACTCGGGGCTCCCTTGGCCCCAGGTGGGCTCCCTTGTCTCTCTGCTCGCT 196
 61 GGGGCGGCTGGGCGGCGGCTTAACTCCCGGACCCCAAGTTCGAGAGCAAGCGGC 120
 197 GGGGCGGCTGGGCGGCGGCTTAACTCCCGGACCCCAAGTTCGAGAGCAAGCGGC 256
 121 TTGCTGGCGCGCGGGGCGCCGGAAGAGCTTCTGTGCTTTCACGAGCGGTTGGAGGACTTG 180

Db 1337 TCAGAGCATCTCTCTGCTCATCTGCTTTGGCTCGAAGCCAGCCAGGAGGAGCCTCT 1396
Qy 1261 GCTGCCAGCTTTGAGTACACTATCTCGGACCCAGCTCCAGCTCTTGCTCCATGGACA 1320
Db 1397 GCTGCCAGCTTTGAGTACACTATCTCGGACCCAGCTCCAGCTCTTGCTCCATGGACA 1456
Qy 1321 CTGTGCCCTGAGTGCCTCCCTACCCACCCACCTAAAGTACCTGTGATCT 1380
Db 1457 CTGTGCCCTGAGTGCCTCCCTACCCACCCACCTAAAGTACCTGTGATCT 1516
Qy 1381 GACTTGGCATCTCACTACACTACAGCTCAGGGACTCCAGGAGCCCAAGGGGCTTA 1440
Db 1517 GACTTGGCATCTCACTACACTACAGCTCAGGGACTCCAGGAGCCCAAGGGGCTTA 1576
Qy 1441 TCGATGGCCCTTACTCCAACTTATGAGACAGCTTATCCAGCCGCTCAGCTCTG 1500
Db 1577 TCGATGGCCCTTACTCCAACTTATGAGACAGCTTATCCAGCCGCTCAGCTCTG 1636
Qy 1501 CCCCCAGCTATGTGGCTTGCTTAG 1527
Db 1637 CCCCCAGCTATGTGGCTTGCTTAG 1663

RESULT 3

ADO05723

ID ADO05723 standard; DNA; 1849 BP.

XX ADO05723;

XX ADO05723;

XX ADO05723;

XX 15-JUL-2004 (first entry)

XX Human erythropoietin receptor (EPOR) encoding DNA.

XX T cell; antiallergic; immunosuppressive; virucide; antibacterial;

KW antiparasitic; cytosolic; gene therapy; human; gene; ds;

KW erythropoietin receptor; EPOR.

XX Homo sapiens.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 137..1663

FT /*tag= a

FT /product= "erythropoietin receptor"

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

XX WO2004032867-A2.

CC benefits from modulating the balance of regulatory T cell function
CC relative to effector T cell function, or vice versa, in a subject. The
CC method involves administering an agent that modulates the expression or
CC activity of a molecule selected from PGER2 and TGFbeta1, or Jagged-1,
CC GPR-32, CD83, CD84, CD89, serotonin R, BY55, serotonin R2C, GPR33, and PI-
CC histamine R-H4, GPR58, EPO-R, PS-1, PS-3, PS-6, PS-9, PS-4d, and PI-
CC 3-related kinase to the subject, such that treatment occurs. The methods
CC are useful for diagnosing, preventing or treating conditions
CC characterized by a too-vigorous or weak effector T cell or regulatory T
CC cell response to antigens associated with the condition, such as in an
CC allergic response, an autoimmune disorder, a viral infection, a microbial
CC infection, a parasitic infection or a tumour. The present sequence
CC represents a DNA encoding a human erythropoietin receptor (EPOR),
XX preferentially expressed in regulatory T cells.

SQ Sequence 1849 BP; 313 A; 593 C; 578 G; 365 T; 0 U; 0 Other;

Query Match 100.0%; Score 1527; DB 12; Length 1849;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGACACACCTCGGGGCTCCCTCTGCGCCCGAGCTCGGCTCCCTTTGTCTCTGCTCGCT 60
Db 137 ATGACACACCTCGGGGCTCCCTCTGCGCCCGAGCTCGGCTCCCTTTGTCTCTGCTCGCT 196
Qy 61 GGGGCCCTGCGGCGCCCGCCCTAACCTCCCGACCCCAAGTTTCGAGAGCAAGCGGCC 120
Db 197 GGGGCCCTGCGGCGCCCGCCCTAACCTCCCGACCCCAAGTTTCGAGAGCAAGCGGCC 256
Qy 121 TTGCTGCGGCGCCCGGCGCGGAGAGCTTCTGTCTTACCCGAGCGGTGGAGACTTG 180
Db 257 TTGCTGCGGCGCCCGGCGCGGAGAGCTTCTGTCTTACCCGAGCGGTGGAGACTTG 316
Qy 181 GTGTGTTTTCGGAGAGAGCGGCGGAGCGCTGGGCTGGCGCCGCACTACAGCTTCTCC 240
Db 317 GTGTGTTTTCGGAGAGAGCGGCGGAGCGCTGGGCTGGCGCCGCACTACAGCTTCTCC 376
Qy 241 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCTTGACACAGGCTTCCACGGCTGT 300
Db 377 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTCGCTTGACACAGGCTTCCACGGCTGT 436
Qy 301 GGTGCGGTGCGCTTCTGCTGCTTACAGCGGAGCGCTGAGCTGCTGCTGCTGCTGCTGCT 360
Db 437 GGTGCGGTGCGCTTCTGCTGCTTACAGCGGAGCGCTGAGCTGCTGCTGCTGCTGCTGCT 496
Qy 361 GAGTTGCGGCTCACAGAGCGCTCCGCGGCTCCCGGATATCACCGTGTCTCATCATCAAT 420
Db 497 GAGTTGCGGCTCACAGAGCGCTCCGCGGCTCCCGGATATCACCGTGTCTCATCATCAAT 556
Qy 421 GAAGTAGTGTCTTAGACGCGCCCGTGGGCTGGTGGCGGCTGGTGGCTGAGAGCGGC 480
Db 557 GAAGTAGTGTCTTAGACGCGCCCGTGGGCTGGTGGCGGCTGGTGGCTGAGAGCGGC 616
Qy 481 CAGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 617 CAGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
Qy 541 GAGTGAGAGCTTCTCGCGCGCAACGCGCGAGGAGCGTACAGAGGTTGGAGATCCTGGAG 600
Db 677 GAGTGAGAGCTTCTCGCGCGCAACGCGCGAGGAGCGTACAGAGGTTGGAGATCCTGGAG 736
Qy 601 GGCGCGACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGAGCGGCTACACCTTGGCGGTC 660
Db 737 GGCGCGACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGAGCGGCTACACCTTGGCGGTC 796
Qy 561 CGCGCGGTATGCTGAGCGGCTTCTGCGGCTTCTGGAGCGCTTGTGGAGCGCTGTG 720
Db 797 CGCGCGGTATGCTGAGCGGCTTCTGCGGCTTCTGGAGCGCTTGTGGAGCGCTGTG 856
Qy 721 TCGCTGTGAGCGCTAGCGACCTGGACCCCTCATCTCTGAGCGCTTCTCCCTCATCTCTGT 780
Db 857 TCGCTGTGAGCGCTAGCGACCTGGACCCCTCATCTCTGAGCGCTTCTCCCTCATCTCTGT 916

PT Treating a condition that benefits from modulating regulatory or effector
PT T cell function comprises administering an agent that modulates the
PT expression or activity of a gene or polypeptide (e.g. PGER2, TGFbeta1,
PT Jagged-1 or GPR-32).

PS Example; SEQ ID NO 23; 161pp; English.

XX The invention relates to treating a subject having a condition that

QY 781 GTGATCTGGTGTCTGTGACCGTGTGCGGTCTCTCCACCGCGGCTCTGAAGCAG 840
 Db 917 GTGATCTGGTGTCTGTGACCGTGTGCGGTCTCTCCACCGCGGCTCTGAAGCAG 976
 QY 841 AAGATCTGGCTGGATCCGAGCCGAGAGCGAGTGTGAGGCTCTTACACACCCAC 900
 Db 977 AAGATCTGGCTGGATCCGAGCCGAGAGCGAGTGTGAGGCTCTTACACACCCAC 1036
 QY 901 AAGGTAATCTTCCAGCTGTGGCTGTACCAAGATGATGCTGTGGTGGAGCCCTTGC 960
 Db 1037 AAGGTAATCTTCCAGCTGTGGCTGTACCAAGATGATGCTGTGGTGGAGCCCTTGC 1096
 QY 961 ACCCCCTTACAGAGAGACCCACTCTCTCCCTGGAGTCTCTACAGAGCTGTGGGG 1020
 Db 1097 ACCCCCTTACAGAGAGACCCACTCTCTCCCTGGAGTCTCTACAGAGCTGTGGGG 1156
 QY 1021 ACGATGAGGAGTGTGGAGCCGGGACAGATGATGAGGCGCCCTCTGAGCCAGTGGC 1080
 Db 1157 ACGATGAGGAGTGTGGAGCCGGGACAGATGATGAGGCGCCCTCTGAGCCAGTGGC 1216
 QY 1081 AGTGAGCATGCCAGATACCTATCTGTGTGGACAAATGGTGTCTCCCGGAAACCCG 1140
 Db 1217 AGTGAGCATGCCAGATACCTATCTGTGTGGACAAATGGTGTCTCCCGGAAACCCG 1276
 QY 1141 CCCAGTGAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGC 1200
 Db 1277 CCCAGTGAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGC 1336
 QY 1201 TCAGAGCATCTCTGCTCATCTGCTTGGCTCGAGCCGAGCCAGCCAGGAGGCTCT 1260
 Db 1337 TCAGAGCATCTCTGCTCATCTGCTTGGCTCGAGCCGAGCCAGGAGGAGCTCT 1396
 QY 1261 GCTGCAGCTTGTAGTACACTATCTGAGCCAGCTCCAGCTCTTGGCTCCATGGACA 1320
 Db 1397 GCTGCAGCTTGTAGTACACTATCTGAGCCAGCTCCAGCTCTTGGCTCCATGGACA 1456
 QY 1321 CTGTGCTGTAGCTGCTCCCTTACCCACCCACCTTAAAGTACCTGTACTTGTGTATCT 1380
 Db 1457 CTGTGCTGTAGCTGCTCCCTTACCCACCCACCTTAAAGTACCTGTACTTGTGTATCT 1516
 QY 1381 GACTCTGGCATCTCAACTGACTTACAGCTCAGGGGACTCCAGGAGCCCAAGGGGGCTTA 1440
 Db 1517 GACTCTGGCATCTCAACTGACTTACAGCTCAGGGGACTCCAGGAGCCCAAGGGGGCTTA 1576
 QY 1441 TCCGATGCCCCCTACTCCAAACCTTATGAGAACAGCCCTTATCCAGCCGCTGAGCCTTG 1500
 Db 1577 TCCGATGCCCCCTACTCCAAACCTTATGAGAACAGCCCTTATCCAGCCGCTGAGCCTTG 1636
 QY 1501 CCCCCAGCTATGGCTGTGCTTAG 1527
 Db 1637 CCCCCAGCTATGGCTGTGCTTAG 1663

RESULT 4

ADP10351
 ID ADP10351 standard; DNA; 1849 BP.

AC ADP10351;
 XX ADP10351;

DT 12-AUG-2004 (first entry)

DE Reference mRNA sequences for marker probe #28.

XX transplant rejection; immune system; rheumatoid arthritis; lupus;
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; SS.

XX Homo sapiens.

XX WO2004042346-A2.

XX 21-MAY-2004.

PD 24-APR-2003; 2003WO-US012946.

XX 24-APR-2002; 2002US-00131831.
 PR 20-DEC-2002; 2002US-00325899.
 XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
 PA Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
 PI Rosenberg S;
 XX WPI; 2004-400724/37.
 DR
 XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
 PT rejection, in an individual, comprises detecting the expression level of
 PT the genes.
 XX
 PS Claim 80; SEQ ID NO 360; 1762pp; English.
 XX
 CC The present invention relates to diagnosing or monitoring transplant
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
 CC comprising detecting the expression level of one or more genes. The
 CC methods, system and kits are useful in diagnosing or monitoring
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
 CC islet, lung, bone marrow or stem cell transplant rejection.
 CC xenotransplant rejection or mechanical organ replacement rejection, in an
 CC individual. The method is also useful in assessing the immune status of
 CC an individual. The methods are also useful in diagnosing and monitoring
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
 CC viral, bacterial or fungal infection. The present sequence represents a
 CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
 CC and monitoring of allograft rejection and other disorders.
 XX
 SQ Sequence 1849 BP; 313 A; 593 C; 578 G; 365 T; 0 U; 0 Other;

Query Match 100.0%; Score 1527; DB 12; Length 1849;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACACCTCGGGGCGTCCCTCTGCCCCAGAGTCGGCTCCCTTTGTCTCTCGCTCGCT 60
 Db 137 ATGACACACCTCGGGGCGTCCCTCTGCCCCAGAGTCGGCTCCCTTTGTCTCTCGCTCGCT 196
 QY 61 GGGGCGCCCTGGGCGCCCCCGCTTAACCTCCCGAACCCCAAGTTCGAGAGCAAGCGGCC 120
 Db 197 GGGGCGCCCTGGGCGCCCCCGCTTAACCTCCCGAACCCCAAGTTCGAGAGCAAGCGGCC 256
 QY 121 TTGCTGGCGCGCGGGGCGCGAGAGCTTCTGTCTTACCGAGCGGTGGAGACTTG 180
 Db 257 TTGCTGGCGCGCGGGGCGCGAGAGCTTCTGTCTTACCGAGCGGTGGAGACTTG 316
 QY 181 GTGTGTTTCTGGGAGGAAGCGGAGCGCTGGGGTGGCGCCGGGCAACTACAGCTTCTCC 240
 Db 317 GTGTGTTTCTGGGAGGAAGCGGAGCGCTGGGGTGGCGCCGGGCAACTACAGCTTCTCC 376
 QY 241 TACAGCTCAGGATGAGCCATGGAAGCTGTGCGCTGACCAAGGCTCCACCGGCTCGT 300
 Db 377 TACAGCTCAGGATGAGCCATGGAAGCTGTGCGCTGACCAAGGCTCCACCGGCTCGT 436
 QY 301 GGTGCGGTGCGCTTCTGTTCTGCTTACAGCCGACACCTCGAGCTTCGTGCCCCCTA 360
 Db 437 GGTGCGGTGCGCTTCTGTTCTGCTTACAGCCGACACCTCGAGCTTCGTGCCCCCTA 496
 QY 361 GAGTTGGCGCTCAGCAGCGCTCCGGGCTCCCGGATATCACCGTGTCAATCAATCAAT 420
 Db 497 GAGTTGGCGCTCAGCAGCGCTCCGGGCTCCCGGATATCACCGTGTCAATCAATCAAT 556
 QY 421 GAAGTAGTGTCTCTAGACGCCCGCTGGGGCTGTGGCGGCTGTGGCTGACGAGCGGC 480
 Db 557 GAAGTAGTGTCTCTAGACGCCCGCTGGGGCTGTGGCGGCTGTGGCTGACGAGCGGC 616
 QY 481 CAGTAGTGTGCTGCTGCTCCCGCTGAGACACCCATGAGCTCTACATCCGCTAC 540

197	GGGGCGGCTTGGGCGCCCGCGGCTTAACCTCCGGGACCCCAAGTTTCGAGAGCAAAAGCGGCC	255
121	TTGCTGGCGGCCGGGGGCCCGGAAGAGCTTCTGTGCTTTCAACGAGCGGTTGGAGACTTG	180
257	TTGCTGGCGGCCGGGGGCCCGGAAGAGCTTCTGTGCTTACCAGAGCGTTGGAGGACTTG	316
181	GTGTGTTTCTGGGAGGAAGCGCGAGCGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC	240
317	GTGTGTTTCTGGGAGGAAGCGCGAGCGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC	376
241	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCTGCAACGAGGCTCCACAGCGTCGT	300
377	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCGCTGCAACGAGGCTCCACAGCGTCGT	436
301	GGTGGGTGCGCTTCTGTGTTTCGTGCTACAGCCGACACGTCGAGCTTCGTGCCCCCTA	360
437	GGTGGGTGCGCTTCTGTGTTTCGTGCTACAGCCGACACGTCGAGCTTCGTGCCCCCTA	496
361	GAGTTGGCGGTCACAGCAGCGCTCCGGCGCTCCGGGATATACGCTGTTCATCCATCAAT	420
497	GAGTTGGCGGTCACAGCAGCGCTCCGGCGCTCCGGGATATACGCTGTTCATCCATCAAT	556
421	GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCTGTGGCGCGGTTGGCTGACGAGAGCGGC	480
557	GAAGTAGTGCTCCTAGACGCGCCCGTGGGGCTGTGGCGCGGTTGGCTGACGAGAGCGGC	616
481	CACGTAGTGTGCGTGCGTCCCGCGCTCAGACACCCATGAGCTCTCAATCCGGTAC	540
617	CACGTAGTGTGCGTGCGTCCCGCGCTCAGACACCCATGAGCTCTCAATCCGGTAC	676
541	GAGTGAGCAGTCTCGGCGGGCAAACGCGGCAAGGAGCGTTACAGAGGTGGAGATCCTGGAG	600
677	GAGTGAGCAGTCTCGGCGGGCAAACGCGGCAAGGAGCGTTACAGAGGTGGAGATCCTGGAG	736
601	GGCGCACCGAGTGTGCTGCTAGCAAACTGCGGGGCGGACGCGCTACACTTCGCGCTC	660
737	GGCGCACCGAGTGTGCTGCTAGCAAACTGCGGGGCGGACGCGCTACACTTCGCGCTC	796
661	CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAGCGCTTGGTCCGAGCCTGTG	720
797	CGCGCGGTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAGCGCTTGGTCCGAGCCTGTG	856
721	TCGCTGTGACGCTTAGGACCTGGACCCCTCATTCCTGAGCGTCTCCCTCATCCTCGTG	780
857	TCGCTGTGACGCTTAGGACCTGGACCCCTCATTCCTGAGCGTCTCCCTCATCCTCGTG	916
781	GTCACTCTGTGTGCTGACGCGTCTCGCGTCTCTCCACCGCGGGTCTCAAGCAG	840
917	GTCACTCTGTGTGCTGACGCGTCTCGCGTCTCTCCACCGCGGGTCTCAAGCAG	976
841	AAGATCTGGCTGGCATCCCGAGCCGAGACGAGCGAGTTGAAGGCTCTTCAACCAACCAC	900
977	AAGATCTGGCTGGCATCCCGAGCCGAGACGAGCGAGTTGAAGGCTCTTCAACCAACCAC	1036
901	AAGGGTAACCTTCAGCTGTGCTGTACAGAAATGATGGTGCCTGTGGTGGAGCCCTTGC	960
1037	AAGGGTAACCTTCAGCTGTGCTGTACAGAAATGATGGTGCCTGTGGTGGAGCCCTTGC	1096
961	ACCCCTTTCAGGAGGACCAACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGTGGGGG	1020
1097	ACCCCTTTCAGGAGGACCAACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGTGGGGG	1156
1021	ACGATGACGACGATGGAGCGGGGACAGATGATGAGGGCCCGCTGCTGGAGCCAGTGGGC	1080
1157	ACGATGACGACGATGGAGCGGGGACAGATGATGAGGGCCCGCTGCTGGAGCCAGTGGGC	1216
1081	AGTGAGCATGCCAGGATACCTATCTGGTGTGGAACAAATGGTTGCTGCCCGGAACCCG	1140
1217	AGTGAGCATGCCAGGATACCTATCTGGTGTGGAACAAATGGTTGCTGCCCGGAACCCG	1276
1141	CCAGTGAGGACCTCCAGGCGCTGTGGACATGTGGACATAGTGGCCATGTGATGAAGGC	1200
1277	CCAGTGAGGACCTCCAGGCGCTGTGGAGTGTGACATAGTGGCCATGTGATGAAGGC	1336

RESULT	6
AAQ05748	
ID	AAQ
XX	AC
AC	XX
DT	25-11-04
DT	04-11-04
XX	XX
DE	EPO
XX	XX
KW	Eryt
OS	Homo
XX	XX
FH	CDS
FT	FT
XX	XX
PN	WO99
PD	09-11-03
XX	XX
PF	03-11-03
XX	XX
PR	(GEN)
XX	(WHI)
PA	Dan
XX	XX
P1	WPI
XX	P-P
DR	Eryt
DR	syst
XX	XX
PS	Clas
XX	XX
CC	The
CC	ava
CC	The
CC	for
CC	the
CC	eryt
CC	ind

CC - cythemia vera. See also AA005747 (murine EPO receptor clone). (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 1883 BP; 320 A; 607 C; 584 G; 372 T; 0 U; 0 Other;

Query Match 100.0%; Score 1527; DB 2; Length 1883; Best Local Similarity 100.0%; Pred. No. 0; Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGACCACTCGGGCGCTCCCTCTGGCCCAAGTCCGCTCCCTTTGCTCTCTGCTCGCT	60
DB	145	ATGGACCACTCGGGCGCTCCCTCTGGCCCAAGTCCGCTCCCTTTGCTCTCTGCTCGCT	204
QY	61	GGGGCGCGCTGGCGCCCGCCCGCTAACTCCCGGACCCCAAGTTCGAGAGCAAAAGCGGCC	120
DB	205	GGGGCGCGCTGGCGCCCGCCCGCTAACTCCCGGACCCCAAGTTCGAGAGCAAAAGCGGCC	264
QY	121	TTGCTGGCGCGCCCGGGGCGCGAAGAGCTTCTGTGCTTCAACGAGCGGTTGGAGACTTG	180
DB	265	TTGCTGGCGCGCCCGGGGCGCGAAGAGCTTCTGTGCTTCAACGAGCGGTTGGAGACTTG	324
QY	181	GTGCTTTTCTGGAGAGAGCGGCGAGCGCTGGGTGGCGCGCAACTACAGCTTCTCC	240
DB	325	GTGCTTTTCTGGAGAGAGCGGCGAGCGCTGGGTGGCGCGCAACTACAGCTTCTCC	384
QY	241	TACAGCTCGAGATGAGCATGGAAGCTGTGCTGCTGACACAGGCTCCACAGGCTCGT	300
DB	385	TACAGCTCGAGATGAGCATGGAAGCTGTGCTGCTGACACAGGCTCCACAGGCTCGT	444
QY	301	GGTGGGCTGGCTTCTGTGTTGCTGCTTACGCGGACAGCTGAGCTTCTGCGCCCTA	360
DB	445	GGTGGGCTGGCTTCTGTGTTGCTGCTTACGCGGACAGCTGAGCTTCTGCGCCCTA	504
QY	361	GAGTTGGGCTCAGAGAGCTTCCGGCGCTCGCGATATACCGGTCTACCATCAAT	420
DB	505	GAGTTGGGCTCAGAGAGCTTCCGGCGCTCGCGATATACCGGTCTACCATCAAT	564
QY	421	GAACTAGTGTCTTAGAGCGCCCGCTGGGGCTGGTGGCGCGGTTGGCTGACGAGCGGC	480
DB	565	GAACTAGTGTCTTAGAGCGCCCGCTGGGGCTGGTGGCGCGGTTGGCTGACGAGCGGC	624
QY	481	CACGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	540
DB	625	CACGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	684
QY	541	GAGTGGAGCTTCTGGCGGCGACAGCGGCGAGGAGCTACAGAGGTTGGAGTCTGGAG	600
DB	685	GAGTGGAGCTTCTGGCGGCGACAGCGGCGAGGAGCTACAGAGGTTGGAGTCTGGAG	744
QY	601	GGCGCACCGAGTGTGCTGAGCAACTGGCGGCGCGGAGCGCTACACCTTGGCGGTC	660
DB	745	GGCGCACCGAGTGTGCTGAGCAACTGGCGGCGCGGAGCGCTACACCTTGGCGGTC	804
QY	661	CGCGCGCTATGCTGAGCGAGCTTCCGGCGCTTCTGGAGCGCTTCTGGAGCGCTGCTG	720
DB	805	CGCGCGCTATGCTGAGCGAGCTTCCGGCGCTTCTGGAGCGCTTCTGGAGCGCTGCTG	864
QY	721	TGCTGCTGAGCGCTGAGCGAAGTGGAGCGGCTTCTGAGCGCTTCTGAGCGCTTCTG	780
DB	865	TGCTGCTGAGCGCTGAGCGAAGTGGAGCGGCTTCTGAGCGCTTCTGAGCGCTTCTG	924
QY	781	GTCATCTGCTGCTGCTGAGCGGCTTCTGAGCGGCTTCTGAGCGGCTTCTGAGCGGCT	840
DB	925	GTCATCTGCTGCTGCTGAGCGGCTTCTGAGCGGCTTCTGAGCGGCTTCTGAGCGGCT	984
QY	841	AAGATCTGGCTGGCATCCCGAGCGCCAGAGAGCGAGTTTGAAGGCTTCTCACCAAGG	900
DB	985	AAGATCTGGCTGGCATCCCGAGCGCCAGAGAGCGAGTTTGAAGGCTTCTCACCAAGG	1044
QY	901	AAGGTAATCTTCCAGCTGTGCTGTACCAAGATGATGCTGCTGCTGCTGCTGCTGCTG	960
DB	1045	AAGGTAATCTTCCAGCTGTGCTGTACCAAGATGATGCTGCTGCTGCTGCTGCTGCTG	1104

QY	961	ACCCCTTTACGGAGGACCCACCTGCTTCCCTGGAAGTCTCTCTCAGAGCGCTGCTGGGG	1020
DB	1105	ACCCCTTTACGGAGGACCCACCTGCTTCCCTGGAAGTCTCTCTCAGAGCGCTGCTGGGG	1164
QY	1021	ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGGAGCCAGTGGGC	1080
DB	1165	ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGGAGCCAGTGGGC	1224
QY	1081	AGTGAGCATCCCGAGGATACCTATCTGCTGCTGACAAATGGTTGCTGCCCGAACCCG	1140
DB	1225	AGTGAGCATCCCGAGGATACCTATCTGCTGCTGACAAATGGTTGCTGCCCGAACCCG	1284
QY	1141	CCAGTGAGGACCTCCAGGCGCTTGGGAGTGTGGACATAGTGGCCATGGATGAAGGC	1200
DB	1285	CCAGTGAGGACCTCCAGGCGCTTGGGAGTGTGGACATAGTGGCCATGGATGAAGGC	1344
QY	1201	TCAGAACATCTCTGCTCATCTGCTTGGCTTGGAGCCAGCCAGAGGAGGCTCT	1260
DB	1345	TCAGAACATCTCTGCTCATCTGCTTGGCTTGGAGCCAGCCAGAGGAGGCTCT	1404
QY	1261	GCTGCCAGCTTTGAGTACACTATCTGAGACCCAGCTCCAGCTCTTGGCTCCATGGACA	1320
DB	1405	GCTGCCAGCTTTGAGTACACTATCTGAGACCCAGCTCCAGCTCTTGGCTCCATGGACA	1464
QY	1321	CTGTGCTTGGCTGCTTACCCACCCAGCTTAAAGTACCTGTACCTTGTGTATCT	1380
DB	1465	CTGTGCTTGGCTGCTTACCCACCCAGCTTAAAGTACCTGTACCTTGTGTATCT	1524
QY	1381	GACTCTGGCATCTCACTGACTACGCTCAGGAGGACTCCAGGAGGAGGAGGCTT	1440
DB	1525	GACTCTGGCATCTCACTGACTACGCTCAGGAGGACTCCAGGAGGAGGAGGCTT	1584
QY	1441	TCCGATGGCGCTTACCTCAACCTTATGAGAAAGAGCTTATCCAGCGCTGAGCCTCTG	1500
DB	1585	TCCGATGGCGCTTACCTCAACCTTATGAGAAAGAGCTTATCCAGCGCTGAGCCTCTG	1644
QY	1501	CCCCCAGCTATGCTGCTCTTCTT	1527
DB	1645	CCCCCAGCTATGCTGCTCTTCTT	1671

RESULT 7
ABL51519
ID ABL51519 standard; cDNA; 1527 BP.
XX ABL51519;
AC ABL51519;
XX 01-JUL-2002 (first entry)
XX Human erythropoietin receptor encoding cDNA SEQ ID NO:4.
XX Human erythropoietin receptor; EPO receptor; fusion protein;
XX erythropoietin receptor extracellular domain; signal transduction;
XX receptor; gene; db.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..1527
XX FT /*tag= a
XX FT /product= "erythropoietin receptor"
XX PN US2002031806-A1.
XX PD 14-MAR-2002.
XX 30-JAN-1998; 98US-00016159.
XX 16-JUN-1997; 97US-00876227.
XX (LEEJ/) LEE J Y.

QY 1141 CCCAGTGGAGCCTCCACGGGCTGTGTCAGTGTGGACATAGTGGCCATGATGAAGGC 1200
 DB 1246 CCCAGTGGAGCCTCCACGGGCTGTGTCAGTGTGGACATAGTGGCCATGATGAAGGC 1305
 QY 1201 TCAGAGGATCCTCTGCTCATCTGCTTTGGGCTCGAAGCCAGCCAGAGGAGGCTCT 1260
 DB 1306 TCAGAGGATCCTCTGCTCATCTGCTTTGGGCTCGAAGCCAGCCAGAGGAGGCTCT 1365
 QY 1261 GCTGCAGCTTGTAGTACATCTCTGGACCCAGCTCCAGCTCTTGGTCCATGGACA 1320
 DB 1366 GCTGCAGCTTGTAGTACATCTCTGGACCCAGCTCCAGCTCTTGGTCCATGGACA 1425
 QY 1321 CTGTGCTGAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTGTGTATCT 1380
 DB 1426 CTGTGCTGAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTGTGTATCT 1485
 QY 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1440
 DB 1486 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1545
 QY 1441 TCCGATGCCCCCTACTCCAAACCTTTATGAGAACAGCTTTATCCAGCCGCTGAGCCTCTG 1500
 DB 1546 TCCGATGCCCCCTACTCCAAACCTTTATGAGAACAGCTTTATCCAGCCGCTGAGCCTCTG 1605
 QY 1501 CCCCCAGCTATGGGCTGCTCTTAG 1527
 DB 1606 CCCCCAGCTATGGGCTGCTCTTAG 1632

RESULT 9

ID ADE28676 standard; cDNA; 1585 BP.
 AC ADE28676;
 XX 29-JAN-2004 (first entry)
 DE Human NOV15d cDNA - SEQ ID 53.
 KW NOVX; antidiabetic; anorectic; cardiatic; hypotensive;
 KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
 KW neurotropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW osteopathic; antiarthritic; antiinflammatory; dermatological;
 KW antiaesthetic; antilipemic; metabolic; diabetes; obesity; infectious;
 KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
 KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
 KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyalipidaemia;
 KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
 KW wound healing; angiogenesis; gene therapy; chromosome mapping;
 KW tissue typing; human; ss; gene; NOV.

OS Homo sapiens.

XX WO2003040330-A2.

PN 15-MAY-2003.

PD 05-NOV-2002; 2002WO-US035536.

PF 05-NOV-2001; 2001US-0338626P.

XX 03-DEC-2001; 2001US-0336600P.

PR 07-DEC-2001; 2001US-0338285P.

PR 12-DEC-2001; 2001US-0341346P.

PR 17-DEC-2001; 2001US-0341477P.

PR 17-DEC-2001; 2001US-0341540P.

PR 20-DEC-2001; 2001US-0342522P.

PR 27-DEC-2001; 2001US-0344297P.

PR 31-DEC-2001; 2001US-0344903P.

PR 17-APR-2002; 2002US-0373288P.

PR 15-MAY-2002; 2002US-0380981P.

PR 17-MAY-2002; 2002US-0381495P.

PR 28-MAY-2002; 2002US-0383534P.

PR 28-MAY-2002; 2002US-0383744P.

29-MAY-2002; 2002US-0383829P.
 29-MAY-2002; 2002US-0384024P.
 07-AUG-2002; 2002US-0401788P.
 26-AUG-2002; 2002US-0406353P.
 31-OCT-2002; 2002US-00287971.
 XX (CURA-) CURAGEN CORP.
 PA Albrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A, Ellerman K, Etenberg S, Gangolli EA, Garlach VL, Gorman L, Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV, Lepley DM, Li L, MacDougall JR, Malyankar UM, Mazur A, McQueney K, Mezes PS, Miller CB, Millet I, Mishra VS, Padigaru M, Patturajan M, Pena CE, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA, Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twonlow N, Vernet CAM, Zerhusen BD, Zhong M;
 WPI; 2003-441555/41.
 P-PSDB; ADE28677.
 New isolated NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
 Claim 20; SEQ ID NO 53; 447pp; English.

The invention relates to a novel isolated NOVX polypeptide. The polypeptide of the invention demonstrates, antidiabetic, anorectic, cardiatic, hypotensive, antiarteriosclerotic, virucide, antibacterial, fungicide, protozoacide, neurotropic, neuroprotective, antiparkinsonian, anticonvulsant, osteopathic, antiarthritic, antiinflammatory, dermatological, antiaesthetic and antilipemic activities. The polypeptides, nucleic acid molecules and antibodies may be useful for treating or diagnosing diseases including metabolic disorders such as diabetes and obesity, infectious diseases, anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and dyalipidaemia. Furthermore, the nucleic acids and polypeptides may also be used to identify molecules that modulate or inhibit neurogenesis, cell differentiation and proliferation, haemopoiesis, wound healing and angiogenesis, as well as in gene therapy. Finally, the nucleic acids may be used as hybridization probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The current sequence is that of the human NOV cDNA of the invention.

Sequence 1585 BP; 263 A; 531 C; 486 G; 305 T; 0 U; 0 Other;

Query Match 99.6%; Score 1520.6; DB 10; Length 1585;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1523; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGACACACCTCGGGGCGTCCCTCTGGCCCGCAGCTCGGCTCTCTTCTCTCTCGCTCGCT 60
 DB 12 ATGACACACCTCGGGGCGTCCCTCTGGCCCGCAGCTCGGCTCTCTTCTCTCTCGCTCGCT 71
 QY 61 GGGGCGCGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 DB 72 GGGGCGCGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 131
 QY 121 TTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 DB 132 TTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 191
 QY 181 GTGTGTTTCTGGAGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 DB 192 GTGTGTTTCTGGAGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 251
 QY 241 TACCAGCTGAGGATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

PA (ELLE/) ELLERMAN K.
 PA (ETE/) ETENBERG S.
 PA (GANG/) GANGOLLI E A.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (GROS/) GROSSE W M.
 PA (GUOX/) GUO X.
 PA (HACK/) HACKETT C.
 PA (JIWW/) JI W.
 PA (KEKU/) KEKUDA R.
 PA (KHRA/) KHRAMTSOV N V.
 PA (LEPL/) LEFLEY D M.
 PA (LILL/) LI L.
 PA (MACD/) MACDOUGALL J R.
 PA (MALY/) MALYANKAR U M.
 PA (MAZU/) MAZUR A.
 PA (MCQU/) MCQUEENEY K.
 PA (MEZE/) MEZES P S.
 PA (MILL/) MILLER C E.
 PA (MILL/) MILLET I.
 PA (MISH/) MISHRA V.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENA C E A.
 PA (PEYM/) PEYMAN J A.
 PA (RAST/) RASTELLI L.
 PA (RIEG/) RIEGER D K.
 PA (ROTH/) ROTHENBERG M E.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (SMIT/) SMITHSON G.
 PA (SPAD/) SPADERNA S K.
 PA (STAR/) STARLING G.
 PA (SPYT/) SPYTEK K A.
 PA (STON/) STONE D J.
 PA (TCHE/) TCHERNEV V T.
 PA (TWOV/) TWOMLOW N.
 PA (VERN/) VERNET C A M.
 PA (ZERH/) ZERHUSEN B D.
 PA (VOSS/) VOSS E Z.
 PA (ZHON/) ZHONG M.
 PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A, Ellerman K, Ettenberg S, Gangolli EA, Gerlach V, Gorman L, Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV, Lepley DM, Li L, MacDougall JR, Malyankar UM, Mazur A, Mcqueeney K, Mezes PS, Miller CE, Millet I, Mishra V, Padigar M, Patturajan M, Pena CEA, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CM, Zerhusen BD, Voss EZ, Zhong M;
 XX WPI; 2004-355303/33.
 DR P-PSDB; ADM93422.
 XX
 PT Novel isolated NOVX polypeptide useful treating or preventing disorders
 PT or syndromes such as Alzheimer's disease, Parkinson's disease, multiple
 PT sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.
 XX
 XX Claim 20; SEQ ID NO 53; 330pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The NOVX polypeptides and polynucleotides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of the sequences in a mammalian subject, and for
 CC treating or preventing a pathology associated with NOVX. The
 CC polypeptides, polynucleotides and antibodies that bind immunospecifically
 CC to the polypeptides are useful for treating or preventing disorders or
 CC syndromes such as congenital heart defects, cardiomyopathy,
 CC atherosclerosis, hypertension, pulmonary stenosis, scleroderma,
 CC adenocarcinoma, haemophilia, graft-versus-host disease, cancer,
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,

CC multiple sclerosis, diabetes, obesity, bronchial asthma, acquired
 CC immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease,
 CC anorexia and immune disorders. This sequence represents a human NOVX
 CC polynucleotide of the invention. Note: The sequence data for this patent
 CC is also available from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1585 BP; 263 A; 531 C; 486 G; 305 T; 0 U; 0 Other;
 Query Match 99.6%; Score 1520.6; DB 12; Length 1585;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1523; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGGACACCTCGGGCGTCCCTTGGCCCCCAGGTGGGCTCCCTTTCTCTCTCGCTCGCT 60
 DB 12 ATGGACACCTCGGGCGTCCCTTGGCCCCCAGGTGGGCTCCCTTTCTCTCTCGCTCGCT 71
 QY 61 GGGGCGGCTTGGGGGCGGCGGCTTAACCTTCCGGACCCCAAGTTCGAGAGCAAGCGGCC 120
 DB 72 GGGGCGGCTTGGGGGCGGCGGCTTAACCTTCCGGACCCCAAGTTCGAGAGCAAGCGGCC 131
 QY 121 TTGCTGGCGGCCCGGGGCGGCGGCTTCTGTGTTTCAACCGAGCGGTTGGAGACTTG 180
 DB 132 TTGCTGGCGGCCCGGGGCGGCGGCTTCTGTGTTTCAACCGAGCGGTTGGAGACTTG 191
 QY 181 GTGTGTTTCTGGAGGAAGCGCGAGCGTGGGTTGGGCCCGGCAACTACAGCTTCTCC 240
 DB 192 GTGTGTTTCTGGAGGAAGCGCGAGCGTGGGTTGGGCCCGGCAACTACAGCTTCTCC 251
 QY 241 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCCTGCACACAGGCTCCACGGCTCGT 300
 DB 252 TACAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCCTGCACACAGGCTCCACGGCTCGT 311
 QY 301 GGTGCGGTGCGCTTCTGTGTTTCCCTTACAGCGACACAGCTCGAGCTTCTGTGCTTCC 360
 DB 312 GGTGCGGTGCGCTTCTGTGTTTCCCTTACAGCGACACAGCTCGAGCTTCTGTGCTTCC 371
 QY 361 GAGTTGGCGTCAACAGAGCTTCCGCGTCCCGGATATCACCGTGTCACTTCCACATCAAT 420
 DB 372 GAGTTGGCGTCAACAGAGCTTCCGCGTCCCGGATATCACCGTGTCACTTCCACATCAAT 431
 QY 421 GAAGTAGTCTCTAGACGCCCGGCTGGTGGCGGCTTGGTGGCGGCTTGGTGGAGAGCGGC 480
 DB 432 GAAGTAGTCTCTAGACGCCCGGCTGGTGGCGGCTTGGTGGCGGCTTGGTGGAGAGCGGC 491
 QY 481 CAGCTAGTGTGGCTTGGCTTCCCGCGCTGAGACACCCATGACGTCTCACATCCGCTAC 540
 DB 492 CAGCTAGTGTGGCTTGGCTTCCCGCGCTGAGACACCCATGACGTCTCACATCCGCTAC 551
 QY 541 GAGGTGAGAGTCTTCCGCCCGGCAACCGCGAGGAGCGGTACAGAGGTTGGAGATCTTGGAG 600
 DB 552 GCGGTGAGAGTCTTCCGCCCGGCAACCGCGAGGAGCGGTACAGAGGTTGGAGATCTTGGAG 611
 QY 601 GCGCGCACGAGTGTGTGTGAGCAACCTTCCGGGCGGAGCGGCTTACACCTTGGCGGTC 660
 DB 612 GCGCGCACGAGTGTGTGTGAGCAACCTTCCGGGCGGAGCGGCTTACACCTTGGCGGTC 671
 QY 661 GCGCGCGGTATGCTGAGCGCGAGCTTCCGCGGCTTCTGTGAGCGCCTTGGTGGAGCGCTGTG 720
 DB 672 GCGCGCGGTATGCTGAGCGCGAGCTTCCGCGGCTTCTGTGAGCGCCTTGGTGGAGCGCTGTG 731
 QY 721 TCCTCTGACGCTAGCGACCTTGGACCCCTCATCTTGAACGCTTCTCCCTCATCTCTCGTG 780
 DB 732 TCCTCTGACGCTAGCGACCTTGGACCCCTCATCTTGAACGCTTCTCCCTCATCTCTCGTG 791
 QY 781 GTCATCTGGTGTCTGACCGTGTCTCCGCTGTCTTCCACCGCGCGGCTCTGAGAGCAG 840
 DB 792 GTCATCTGGTGTCTGACCGTGTCTCCGCTGTCTTCCACCGCGCGGCTCTGAGAGCAG 851
 QY 841 AAGATCTGCGCTGGCATCCCGAGCGCCAGAGCGAGTGTGAAGCGCTCTTCAACACCCAC 900
 DB 852 AAGATCTGCGCTGGCATCCCGAGCGCCAGAGCGAGTGTGAAGCGCTCTTCAACACCCAC 911
 QY 901 AAGGTAACCTTCCAGCTGTGGCTGTGTACCAAGATGTAGGCTGTGCTGGTGGAGCCCTCTCG 960

912 AAGGTAACCTCCAGCTGTGGCTGTACCAATGATGGCTGCTGTGGTGGAGCCCTGCTGC 971
 961 ACCCCCTTACGAGGAGACCACTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGG 1020
 972 ACCCCCTTACGAGGAGACCACTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGGG 1031
 1021 ACCATGAGGAGCTGGAGCGCGGACAGATGATGAGGGCCCTGCTGGAGCCAGTGGGC 1080
 1032 ACCATGAGGAGCTGGAGCGCGGACAGATGATGAGGGCCCTGCTGGAGCCAGTGGGC 1091
 1081 AGTGAGATGCCAGGATACCTATCTGGTCTGGAATAATGGTGTGCTGCCCGAACCCG 1140
 1092 AGTGAGATGCCAGGATACCTATCTGGTCTGGAATAATGGTGTGCTGCCCGAACCCG 1151
 1141 CCAGTGAGGAGCTCCAGGGCTGTGGCAGTGTGACATAGTGGCCATGATGAAGC 1200
 1152 CCAGTGAGGAGCTCCAGGGCTGTGGCAGTGTGACATAGTGGCCATGATGAAGC 1211
 1201 TCAGAGGATCTCTCTGCTCATCTGCTTTGGCCCTCGAAGCCAGCCAGAGGAGCTCT 1260
 1212 TCAGAGGATCTCTCTGCTCATCTGCTTTGGCCCTCGAAGCCAGCCAGAGGAGCTCT 1271
 1261 GCTGCCAGCTTTGAGTACACTATCTCTGAGACCCAGCTTCCAGCTTTGGTCCATGGACA 1320
 1272 GCTGCCAGCTTTGAGTACACTATCTCTGAGACCCAGCTTCCAGCTTTGGTCCATGGACA 1331
 1321 CTGTGCCCTGAGTGCCTCCCTACCCACCCACCTAAAGTACTGCTGACCTTGGTATCT 1380
 1332 CTGTGCCCTGAGTGCCTCCCTACCCACCCACCTAAAGTACTGCTGACCTTGGTATCT 1391
 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGAGCCCAAGGGGCTTA 1440
 1392 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGAGCCCAAGGGGCTTA 1451
 1441 TCGATGGCCCTTACTTCCAACTTATGAGAACAGCTTATCCAGCCCTGAGCTCTG 1500
 1452 TCGATGGCCCTTACTTCCAACTTATGAGAACAGCTTATCCAGCCCTGAGCTCTG 1511
 1501 CCCCCCAGCTATGTGGCTTGTCTTAG 1527
 1512 CCCCCCAGCTATGTGGCTTGTCTTAG 1538

RESULT 11
 AAQ53995
 ID AAQ53995 standard; cDNA; 1818 BP.
 AC AAQ53995;
 DT 25-MAR-2003 (revised)
 DT 24-JUN-1994 (first entry)
 DE Human EPO receptor cDNA.
 XX
 XX Erythropoietin receptor; recombinant; murine; anaemia; ss.
 XX Homo sapiens.
 Key Location/Qualifiers
 FT CDS 106..1632
 FT /*tag= a
 FT sig_peptide 106..177
 FT /*tag= b
 FT /*note= "signal peptide"
 XX
 PN US278065-A.
 XX
 XX 11-JAN-1994.
 XX
 PF 25-MAR-1991; 91US-00678877.
 XX
 PR 03-FEB-1989; 89US-00306503.

(GEMV) GENETICS INST INC.
 (CHIL-) CHILDRENS MEDICAL CENT.
 (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PI D'andrea A, Wong GG, Jones SS;
 WPI: 1994-025409/03.
 P-PSDB; AAR47518.
 Recombinant DNA encoding erythropoietin receptor - used to develop prods.
 for study, treatment or diagnosis of disorders in which receptor is
 dysfunctional.
 Claim 7; Fig 9; 24pp; English.
 Mouse erythroleukaemia (MEL) cells were used to construct a cDNA library.
 The cDNA was used to transfect COS-1 cells and these were screened for
 radioiodinated erythropoietin (EPO) binding to isolate cDNA encoding the
 EPO receptor. This cDNA was used as a probe to screen a human genomic
 cDNA library to obtain DNA encoding the human EPO receptor. The cDNA may
 be used to study, treat or diagnose disorders in which the EPO receptor
 is dysfunctional. See also AA053994. (Updated on 25-MAR-2003 to correct
 PF field.)
 Sequence 1818 BP; 308 A; 579 C; 360 T; 0 U; 0 Other;
 Query Match 99.3%; Score 1515.8; DB 2; Length 1818;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1520; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGGACCACTCTGGGGCGCTCCCTCTGGCCCGAGGTCGGCTCTCTTGTCTCTGCTCGCT 60
 DB 106 ATGGACCACTCTGGGGCGCTCCCTCTGGCCCGAGTGGCTCTCTTGTCTCTGCTCGCT 165
 QY 61 GGGGCGCGCTGGGGCGCCCGCCCTAACTCCCGGACCCCAAGTTCCAGAGCAAAAGCGGC 120
 DB 166 GGGGCGCGCTGGGGCGCCCGCCCTAACTCCCGGACCCCAAGTTCCAGAGCAAAAGCGGC 225
 QY 121 TTGCTGGCGGCGCGGGGCGCGAGAGCTTCTGTCTTCCAGGCGGTTGGAGACTTG 180
 DB 226 TTGCTGGCGGCGCGGGGCGCGAGAGCTTCTGTCTTCCAGGCGGTTGGAGACTTG 285
 QY 181 GTGTGTTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGCGCCGCAACTACAGCTTCTCC 240
 DB 286 GTGTGTTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGCGCCGCAACTACAGCTTCTCC 345
 QY 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCACAGGCTCCCAAGGCTCGT 300
 DB 346 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCACAGGCTCCCAAGGCTCGT 405
 QY 301 GGTGGGTGGCGCTTCTGTGTTCGCTTACGCGACGCTCGAGCTTGTGCGCCCTA 360
 DB 406 GGTGGGTGGCGCTTCTGTGTTCGCTTACGCGACGCTCGAGCTTGTGCGCCCTA 465
 QY 361 GAGTTGCGCGTCCAGCAGCGCTCCGGCGCTCCGGGATATCACCGTGTATCATCAATCAAT 420
 DB 466 GAGTTGCGCGTCCAGCAGCGCTCCGGCGCTCCGGGATATCACCGTGTATCATCAATCAAT 525
 QY 421 GAAGTAGTGTCTTAGACGCGCGCTGGGGCTGTGGCGCGGTTGGCTGACGAGAGCGGC 480
 DB 526 GAAGTAGTGTCTTAGACGCGCGCTGGGGCTGTGGCGCGGTTGGCTGACGAGAGCGGC 585
 QY 481 CACGTAGTGTGGCTGGCTCCCGCGCTCAGACACCCATGAGCTCTCACATCCGCTAC 540
 DB 586 CACGTAGTGTGGCTGGCTCCCGCGCTCAGACACCCATGAGCTCTCACATCCGCTAC 645
 QY 541 GAGGTGACAGCTCTCGCGCGCAACGCGGCGAGGAGCGGTACAGAGGGTGGAGATCTCGAG 600
 DB 646 GAGGTGACAGCTCTCGCGCGCAACGCGGCGAGGAGCGGTACAGAGGGTGGAGATCTCGAG 705
 QY 601 GGCGCGACCGAGTGTGTCTGAGCAACTTGGGGGCGGAGCGCTACACCTTGGCGCTC 660

```
Db 706 GGGCCGACCGAGTGTGTCTGACGAACTCTGGGGGCGGACGCGCTACACCTTTCGGCGTC 765
Qy 661 CGCGCGGTATGGCTGAGCGAGCTTTCGGCGGCTTCTGGAGCGCTTGGTGGAGCGCTGTG 720
Db 766 CGCGCGGTATGGCTGAGCGAGCTTTCGGCGGCTTCTGGAGCGCTTGGTGGAGCGCTGTG 825
Qy 721 TCGCTGCTGACGCTAGCGACCTGGACCGCCCTCATCTGACGCTTCCCTCATCTCTGCTG 780
Db 826 TCGCTGCTGACGCTAGCGACCTGGACCGCCCTCATCTGACGCTTCCCTCATCTCTGCTG 885
Qy 781 GTCATCTCTGCTGCTGACCGGTCTCGGCTGCTCTCCACCGCGGCTCTGAAGCAG 840
Db 886 GTCATCTCTGCTGCTGACCGGTCTCGGCTGCTCTCCACCGCGGCTCTGAAGCAG 945
Qy 841 AAGATCTGCTGGCATCCCGAGCGCCAGAGAGCGAGTTCGAAGGCTCTTACACCCAC 900
Db 946 AAGATCTGCTGGCATCCCGAGCGCCAGAGAGCGAGTTCGAAGGCTCTTACACCCAC 1005
Qy 901 AAGGTAACCTTCAGCTGTGGCTGTACAGATGATGGCTGCTGTGGTGGAGCGCTCTG 960
Db 1006 AAGGTAACCTTCAGCTGTGGCTGTACAGATGATGGCTGCTGTGGTGGAGCGCTCTG 1065
Qy 961 ACCCCCTTCACGAGGACCCACCTGCTTCCCTGGAGTCTCTCAGAGCGCTCTGGGG 1020
Db 1066 ACCCCCTTCACGAGGACCCACCTGCTTCCCTGGAGTCTCTCAGAGCGCTCTGGGG 1125
Qy 1021 ACCATGAGGAGTGGAGCGGCGGACAGATGATAGGGGCCCCCTGCTGGAGCGAGTGGG 1080
Db 1126 ACGATGAGGAGTGGAGCGGCGGACAGATGATAGGGGCCCCCTGCTGGAGCGAGTGGG 1185
Qy 1081 AGTGAGCATGCCAGGATACCTATCTGTGTGTGGCAAAATGTTGTCGCCCGGAACCG 1140
Db 1186 AGTGAGCATGCCAGGATACCTATCTGTGTGTGGCAAAATGTTGTCGCCCGGAACCG 1245
Qy 1141 CCCAGTGAGGACTCCAGGGCTGTGGCGCTGTGGCATAGTGGCATGATGAGGC 1200
Db 1246 CCCAGTGAGGACTCCAGGGCTGTGGCGCTGTGGCATAGTGGCATGATGAGGC 1305
Qy 1201 TCAGAGCATCTCTGCTCATCTGCTTGGGCTCGAAGCCCGAGCGGAGGCTCT 1260
Db 1306 TCAGAGCATCTCTGCTCATCTGCTTGGGCTCGAAGCCCGAGCGGAGGCTCT 1365
Qy 1261 GCTGCGAGTTGAGTACATCTCTGAGACCCAGCTCCAGCTCTTGGTCCATGGACA 1320
Db 1366 GCTGCGAGTTGAGTACATCTCTGAGACCCAGCTCCAGCTCTTGGTCCATGGACA 1425
Qy 1321 CTGTGCGCTGAGCTGCCCTTACCCACCCACCTAAGTACCTGTACCTTGTGTATCT 1380
Db 1426 CTGTGCGCTGAGCTGCCCTTACCCACCCACCTAAGTACCTGTACCTTGTGTATGT 1485
Qy 1381 GACTCTGGCATCTCACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1440
Db 1486 GACTCTGGCATCTCACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGGCTTA 1545
Qy 1441 TCGATGGCCCTACTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG 1500
Db 1546 TCGATGGCCCTACTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG 1605
Qy 1501 CCCCCCAGTATGTGGCTTGTCTTAG 1527
Db 1606 CCCCCCAGTATGTGGCTTGTCTTAG 1632
```

RESULT 12

AAZ49634

ID AAZ49634 standard; cDNA; 1317 BP.

XX AC

AAZ49634;

XX AC

DT 07-APR-2000 (first entry)

XX AC

DE Truncated human EpoR (t439) cDNA.

XX AC

Truncated human EpoR; erythropoietin receptor; hypersensitive EpoR(t439); mutant human EpoR; EpoR signalling; cancer; infectious disease; HIV; sickle cell anaemia; cytosolic; antimicrobial; antiviral; immunostimulant; anti-anaemic; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1317
/*tag= a
/product= "Truncated human EpoR(t439) "

WO9967360-A2.

29-DEC-1999.

25-JUN-1999; 99WO-CA000606.

25-JUN-1999; 98CA-02241576.

25-JAN-1999; 99CA-02260332.

(HEMO-) HEMOSOL INC.

Bell D, Matthews KE, Mueller SG;

WPI; 2000-136979/12.

P-PSDB; AAY44622.

Serum free defined medium useful for the efficient culture of stem cells used for production of hemoglobin.

Example 6; Fig 9; 61pp; English.

The present cDNA sequence encodes truncated human EpoR (erythropoietin receptor). This was isolated from human umbilical cord blood LDMNC using standard PCR methods. Transfection of constitutively active EpoR(t439) by electroporation into a cytokine-dependent cell line supports cell population expansion in the absence of exogenous cytokines. Mutant human EpoR is used in treatment of disorders related to inadequate EpoR signalling. The transfected cells may also be used in gene therapy to treat cancer, infectious diseases (e.g. HIV), sickle cell anaemia, and conditions related to abnormal expression of erythropoietin.

SQ Sequence 1317 BP; 207 A; 437 C; 422 G; 251 T; 0 U; 0 Other;

Query Match 85.9%; Score 1312.2; DB 3; Length 1317;

Best Local Similarity 99.8%; Pred. No. 2.3e-272;

Matches 1314; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGACCACTCTGGGGGCTCCCTCTGGCCCGAGTGGCTCCCTTTGTCTCTGCTCGCT 60

Db 1 ATGACCACTCTGGGGGCTCCCTCTGGCCCGAGTGGCTCCCTTTGTCTCTGCTCGCT 60

Qy 61 GGGGCCCTGGGGCGCCCCCGCTAACTCCCGACCCCAAGTTCGAGAGCAAGCGGCC 120

Db 61 GGGGCCCTGGGGCGCCCCCGCTAACTCCCGACCCCAAGTTCGAGAGCAAGCGGCC 120

Qy 121 TTGCTGGCGGCGCGGGGCGCCGAAAGAGCTTCTGTCTTACCGAGCGGTTGGAGACTTG 180

Db 121 TTGCTGGCGGCGCGGGGCGCCGAAAGAGCTTCTGTCTTACCGAGCGGTTGGAGACTTG 180

Qy 181 GTGTGTTTCTGGAGAGGCGGCGGCTGGGGTGGGCGCCGGGCAACTACAGCTTCTCC 240

Db 181 GTGTGTTTCTGGAGAGGCGGCGGCTGGGGTGGGCGCCGGGCAACTACAGCTTCTCC 240

Qy 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCTGCTGACCGACAGCTCGAGCTTCC 300

Db 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCTGCTGACCGACAGCTTCC 300

Qy 301 GGTGCGGTGGCTTCTGTGTCTGCTGCTACAGCCGACAGCTCGAGCTTCTGCCCCCTA 360

Db 301 GGTGCGGTGGCTTCTGTGTCTGCTGCTACAGCCGACAGCTCGAGCTTCTGCCCCCTA 360

Db 301 GGTGCGGTACGCTTCTGGTTCGCTGCTACAGCCGACACATCGAGCTTCGTGCCCCCTA 360
 Qy 361 GAGTTGCGGCTCACAGCAGCTTCCGCGCTCCGCGATATCACCGTGTATCCACATCAAT 420
 Db 361 GAGTTGCGGCTCACAGCAGCTTCCGCGCTCCGCGATATCACCGTGTATCCACATCAAT 420
 Qy 421 GAAAGTAGTCTCTAGACGCCCCCGTGGGGCTGGTGGCGGTTGGCTGACAGAGCGGC 480
 Db 421 GAAAGTAGTCTCTAGACGCCCCCGTGGGGCTGGTGGCGGTTGGCTGACAGAGCGGC 480
 Qy 481 CAGCTAGTGTGGCTTCCGCGCTCCGCGCTGAGACACCATCATCGCTCTCACATCCGCTAC 540
 Db 481 CAGCTAGTGTGGCTTCCGCGCTCCGCGCTGAGACACCATCATCGCTCTCACATCCGCTAC 540
 Qy 541 GAGGTGGAGCTTCTCGCGCGGCAACCGCGCAGGAGCGTACAGAGGTTGGAGTCTGGAG 600
 Db 541 GAGGTGGAGCTTCTCGCGCGGCAACCGCGCAGGAGCGTACAGAGGTTGGAGTCTGGAG 600
 Qy 601 GCGCGGACCGAGTGTGTGAGCAACCTGCGGGCGCGGACGCGCTACACCTTCCGCGTC 660
 Db 601 GCGCGGACCGAGTGTGTGAGCAACCTGCGGGCGCGGACGCGCTACACCTTCCGCGTC 660
 Qy 661 CCGCGCGGTATGGCTGAGCGGAGCTTCCGCGGCTTCTGGAGCGCTTGGAGCGCTGTG 720
 Db 661 CCGCGCGGTATGGCTGAGCGGAGCTTCCGCGGCTTCTGGAGCGCTTGGAGCGCTGTG 720
 Qy 721 TCGCTGTGACCGCTAGCAGCTGACACCCCTCATCTCTGACGCTCTCCCTCATCTCTG 780
 Db 721 TCGCTGTGACCGCTAGCAGCTGACACCCCTCATCTCTGACGCTCTCCCTCATCTCTG 780
 Qy 781 GTCATCTGTGCTGTGACCGCTGTGCGCTGTCTTCCACCGCGGCTCTGAAGCAG 840
 Db 781 GTCATCTGTGCTGTGACCGCTGTGCGCTGTCTTCCACCGCGGCTCTGAAGCAG 840
 Qy 841 AGATCTGGCTGATCCCGAGCCGAGAGCGAGTGTGAAGGCTCTTACACACCCAC 900
 Db 841 AGATCTGGCTGATCCCGAGCCGAGAGCGAGTGTGAAGGCTCTTACACACCCAC 900
 Qy 901 AAGGTTAACTTCCAGCTGTGGCTGTACAGAAATGATGCTGTGGTGGAGCCCTGTC 960
 Db 901 AAGGTTAACTTCCAGCTGTGGCTGTACAGAAATGATGCTGTGGTGGAGCCCTGTC 960
 Qy 961 ACCCTTTCACGAGAGCCACCTGCTTCCCTTGGAAATGCTCTCAGAGCGCTGTGGGG 1020
 Db 961 ACCCTTTCACGAGAGCCACCTGCTTCCCTTGGAAATGCTCTCAGAGCGCTGTGGGG 1020
 Qy 1021 ACATGACGAGTGTGGAGCCGGGACAGTATGATGAGGCCCCCTCTGAGCCAGTGGGC 1080
 Db 1021 ACATGACGAGTGTGGAGCCGGGACAGTATGATGAGGCCCCCTCTGAGCCAGTGGGC 1080
 Qy 1081 AGTGAGCATGCCAGGATACCTATCTGGTGTGGACAAATGTTGCTGCCCCGGAACCCG 1140
 Db 1081 AGTGAGCATGCCAGGATACCTATCTGGTGTGGACAAATGTTGCTGCCCCGGAACCCG 1140
 Qy 1141 CCAGTGTAGGACCTCCAGGGCTTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGC 1200
 Db 1141 CCAGTGTAGGACCTCCAGGGCTTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGC 1200
 Qy 1201 TCAGAGCATCTCTGCTCATCTGTTGGCTTGGCCCTGAGCCCGAGCGGAGGCGCTCT 1260
 Db 1201 TCAGAGCATCTCTGCTCATCTGTTGGCTTGGCCCTGAGCCCGAGCGGAGGCGCTCT 1260
 Qy 1261 GCTGCGAGCTTTGAGTACATCTCTGAGCCCGAGCTCCAGCTCTTGGCTCCATGG 1317
 Db 1261 GCTGCGAGCTTTGAGTACATCTCTGAGCCCGAGCTCCAGCTCTTGGCTCCATGG 1317

RESULT 14
 ADE28670
 ID ADE28670 standard; cDNA; 1733 BP.
 XX
 AC ADE28670;
 XX

DT XX 29-JAN-2004 (first entry).
 DE XX Human NOV15a cDNA - SEQ ID 47.
 XX XX
 KW NOVX; antidiabetic; anorectic; cardiant; hypotensive;
 KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;
 KW neotropic; neuroprotective; antiparkinsonian; anticonvulsant;
 KW osteopathic; antiarthritic; antiinflammatory; dermatological;
 KW antiasthmatic; antilipemic; metabolic; diabetes; obesity; infectious;
 KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;
 KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;
 KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;
 KW neurogenesis; cell differentiation; proliferation; haemopoiesis;
 KW wound healing; angiogenesis; gene therapy; chromosome mapping;
 KW tissue typing; human; ss; gene; NOV.
 OS Homo sapiens.
 XX
 PN WO2003040330-A2.
 XX
 PD 15-MAY-2003.
 XX
 PP 05-NOV-2002; 2002WO-US035536.
 XX
 PR 05-NOV-2001; 2001US-0338626P.
 PR 05-DEC-2001; 2001US-0336600P.
 PR 07-DEC-2001; 2001US-0338285P.
 PR 12-DEC-2001; 2001US-0341346P.
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 20-DEC-2001; 2001US-0342522P.
 PR 27-DEC-2001; 2001US-0344297P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383534P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0383829P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 26-AUG-2002; 2002US-0406533P.
 PR 31-OCT-2002; 2002US-00287971.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL;
 PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
 PI Ellerman K, Ettenberg S, Gangoli EA, Garlach VL, Gorman L;
 PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
 PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, Mcqueeny K;
 PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigar M, Patturajan M;
 PI Pena CEA, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;
 PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
 PI Vernet CAM, Zerkusen BD, Zhong M;
 PI WPI; 2003-441555/41.
 DR P-PSDB; ADE28671.
 DR
 XX
 XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 20; SEQ ID NO 47; 447pp; English.
 XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
 CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
 CC fungicide, protozoacide, neurotropic, neuroprotective, antiparkinsonian,
 CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
 CC dermatological, antiasthmatic and antilipemic activities. The
 CC polypeptides, nucleic acid molecules and antibodies may be useful for

treating or diagnosing diseases including metabolic disorders such as diabetes and obesity, infectious diseases, anorexia, cancer, cardiovascular diseases including hypertension and atherosclerosis, neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic disorders, inflammatory skin disorders, asthma and dyslipidaemia. Furthermore, the nucleic acids and polypeptides may also be used to identify molecules that modulate or inhibit neurogenesis, cell differentiation and proliferation, haemopoiesis, wound healing and angiogenesis, as well as in gene therapy. Finally, the nucleic acids may be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The current sequence is that of the human NOV cDNA of the invention.

Sequence 1733 BP; 291 A; 558 C; 546 G; 338 T; 0 U; 0 Other; ;

Query Match	79.2%	Score 1209;	DB 10;	Length 1733;
Best Local Similarity	89.8%	Pred. No. 3.6e-250;		
Matches 1372;	Conservative	0;	Mismatches 5;	Indels 150;
Gaps	1			

1	ATGAGCACCTCGGGCGTCCCTCTGGGCCCCAGTCCGGCTCCCTTTGTTCTCTGCTCGCT	60
145	ATGAGCACCTCGGGCGTCCCTCTGGGCCCCAGTCCGGCTCCCTTTGTTCTCTGCTCGCT	204
61	GGGGCGGCTGGGCGCCCCCGCCTAACTCCGGACCCCAAGTTCGAGAGCAAGCGGCC	120
205	GGGGCGGCTGGGCGCCCCCGCCTAACTCCCGGACCCCAAGTTCGAGAGCAAGCGGCC	264
121	TTTGTGGGGCCCCGGGGCCCCGAGAGAGCTTCTGTGCTTCACTCCGAGCGTTGGAGGACTTG	180
265	TTTGTGGGGCCCCGGGGCCCCGAGAGAGCTTCTGTGCTTCACTCCGAGCGTTGGAGGACTTG	324
181	GTGTGTTTCTGGGAGAAAGCGGAGCGCTGGGCTGGCGCCGGCAACTACAGCTTCTCC	240
325	GTGTGTTTCTGGGAGAAAGCGGCGAGCGCTGGGCTGGCGCCGGCAACTACAGCTTCTCC	384
241	TACCAAGCTCGAGGATGAGCCATGGAAGCTGTGTCCGCTGCACAGAGCTCCACGGTCTCGT	300
385	TACCAAGCTCGAGGATGAGCCATGGAAGCTGTGTCCGCTGCACAGAGCTCCACGGTCTCGT	444
301	GGTGGGTGGCTTCTGGTGTTCGTCATACGCCACAGCTCGAGCTTCGTGGCCCTTA	360
445	GGTGGGTGGCTTCTGGTGTTCGTCGCTACAGCCGACAGCTCGAGCTTCGTGGCCCTTA	504
361	GAGTTTCCGGTACAGAGAGCTCCGGCGCTCCGGGATACACCGTGTCACTCCACATCAAT	420
505	GAGTTTCCGGTACAGAGAGCTCCGGCGCTCCGGATACACCGTGTCACTCCACATCAAT	564
421	GAAGTAGTGTCTCTAGACGCCCGCTGGGGCTGTGTGGCGGTGGCTGCACAGAGCGGC	480
565	GAAGTAGTGTCTCTAGACGCCCGCTGGGGCTGTGTGGCGGTGGCTGCACAGAGCGGC	624
481	CACGTAGTGTTCGCTGGCTCCCGCGCTCGAGACACCCATGAGCTCTCACTCCGCTAC	540
625	CACGTAGTGTTCGCTGGCTCCCGCGCTCGAGACACCCATGAGCTCTCACTCCGCTAC	684
541	GAGTGGAGCTTCTCGGCCGCAACGGCGCAGGAGCGTACAGAGGTGGAGATCTCTGGAG	600
685	GAGTGGAGCTTCTCGGCCGCAACGGCGCAGGAGCGTACAGAGGTGGAGATCTCTGGAG	744
601	GGCCGCAACGAGTGTGTGAGCAACTCTGGGGGGCGGACGGCTACACCTTCGCGCTC	660
745	GGCCGCAACGAGTGTGTGAGCAACTCTGGGGGGCGGACGGCTACACCTTCGCGCTC	804
661	CGCGCGCTATGGCTGAGCGAGCTTTCGGCGGCTTCTGGAGCGCTGTGGAGCGCTGTG	720
805	CGCGCGCTATGGCTGAGCGAGCTTTCGGCGGCTTCTGGAGCGCTGTGGAGCGCTGTG	864
721	TCGCTGCTGAGCCCTAGGCACTGGACCCCTCATCTCTGAGCTCTCCCTCATCTCTCGT	780
865	TCGCTGCTGAGCCCTAGGCACTGGACCCCTCATCTCTGAGCTCTCCCTCATCTCTCGT	924
781	GTCACTCTGGTCTGTGACCGGTGCTCGCGTGTCTCTCCCAACCGCGGGCTCTGAAGACG	840

XX Homo sapiens.
 OS WO2003040330-A2.
 PN 15-MAY-2003.
 XX 05-NOV-2002; 2002WO-US035536.
 XX 05-NOV-2001; 2001US-0338626P.
 XX 05-DEC-2001; 2001US-0336600P.
 PR 07-DEC-2001; 2001US-0338285P.
 PR 12-DEC-2001; 2001US-0341346P.
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 20-DEC-2001; 2001US-0342592P.
 PR 27-DEC-2001; 2001US-0344297P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383534P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0383629P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 26-AUG-2002; 2002US-0406353P.
 PR 31-OCT-2002; 2002US-00287971.
 XX (CURA-) CURAGEN CORP.
 PA
 XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boidog FL;
 PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;
 PI Ellerman K, Ertzenberg S, Gangolli EA, Gerlach VL, Gorman L;
 PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;
 PI Lepley DM, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K;
 PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigaru M, Patturajan M;
 PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;
 PI Smithson G, Starling G, Spytek KA, Stone DJ, Tchernev VT, Twomlow N;
 PI Vernet CAM, Zerhusen BD, Zhong M;
 XX WPI; 2003-441555/41.
 DR P-PSDB; ADE28673.
 XX
 XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.
 XX
 PS Claim 20; SEQ ID NO 49; 447pp; English.
 XX
 CC The invention relates to a novel isolated NOVX polypeptide. The
 CC polypeptide of the invention demonstrates, antidiabetic, anorectic,
 CC cardiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,
 CC fungicide, protozoacide, nootropic, neuroprotective, antiparkinsonian,
 CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,
 CC dermatological, antiasthmatic and antilipemic activities. The
 CC polypeptides, nucleic acid molecules and antibodies may be useful for
 CC treating or diagnosing diseases including metabolic disorders such as
 CC diabetes and obesity, infectious diseases, anorexia, cancer,
 CC cardiovascular diseases including hypertension and atherosclerosis,
 CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's
 CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic
 CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.
 CC Furthermore, the nucleic acids and polypeptides may also be used to
 CC identify molecules that modulate or inhibit neurogenesis, cell
 CC differentiation and proliferation, haemopoiesis, wound healing and
 CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may
 CC be used as hybridisation probes, in chromosome mapping, tissue typing,
 CC preventive medicine and pharmacogenomics. The current sequence is that of
 CC the human NOV CDNA of the invention.
 XX
 SQ Sequence 1733 BP; 291 A; 558 C; 546 G; 338 T; 0 U; 0 Other;

Query Match	79.2%	Score	1209	DB	10	Length	1733
Best Local Similarity	89.8%	Pred. No.	3.6e-250				
Matches	1372	Conservative	0	Mismatches	5	Indels	150
Gaps	1						
QY	1	ATGGACACACCTCGGGGGCTCCCTCTGCGCCAGCTCGGCTCCCTTTGTCTCTGCTCGCT	60				
DB	145	ATGGACACACCTCGGGGGCTCCCTCTGCGCCAGCTCGGCTCCCTTTGTCTCTGCTCGCT	204				
QY	61	GGGGCCGCTTGGGGCCCCCGCTTAACCTCCCGGACCCCAAGTTTCGAGAGCAAAAGCGCC	120				
DB	205	GGGGCCGCTTGGGGCCCCCGCTTAACCTCCCGGACCCCAAGTTTCGAGAGCAAAAGCGCC	264				
QY	121	TTGCTGGCGGCGCGGGGGCCGGAAGAGCTTCTGTGCTTACCGAGCGGTTGAGAGACTTG	180				
DB	265	TTGCTGGCGGCGCGGGGGCCGGAAGAGCTTCTGTGCTTACCGAGCGGTTGAGAGACTTG	324				
QY	181	GTGTGTTTCTGGAGGAAGCGGAGCGCTGGGTTGGGCGCGGCAACTACAGCTTCTTCC	240				
DB	325	GTGTGTTTCTGGAGGAAGCGGAGCGCTGGGTTGGGCGCGGCAACTACAGCTTCTTCC	384				
QY	241	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCCTGCACACAGGCTCCCAAGGCTCGT	300				
DB	385	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTTGCCTGCACACAGGCTCCCAAGGCTCGT	444				
QY	301	GGTGGGTGCGCTTCTGTGTTGCTGCTTACAGCGGACACGTCGAGCTTCTGTGCCCCCTA	360				
DB	445	GGTGGGTGCGCTTCTGTGTTGCTGCTTACAGCGGACACGTCGAGCTTCTGTGCCCCCTA	504				
QY	361	GAGTTGGCGGTCACAGCAGCCTCGGGCGTCCGGGATATCACCGTGTATCCACATCAAT	420				
DB	505	GAGTTGGCGGTCACAGCAGCCTCGGGCGTCCGGGATATCACCGTGTATCCACATCAAT	564				
QY	421	GAAGTAGTGTCTTAGACGCGCCGCTGGGCTGTGTGGCGGTTGGCTGACGAGAGCGCG	480				
DB	565	GAAGTAGTGTCTTAGACGCGCCGCTGGGCTGTGTGGCGGTTGGCTGACGAGAGCGCG	624				
QY	481	CACGTAGTGTGCGCTGGCTTCCGCGCGCTGAGACACCCATGAGCTTCTACATCCGCTAC	540				
DB	625	CACGTAGTGTGCGCTGGCTTCCGCGCGCTGAGACACCCATGAGCTTCTACATCCGCTAC	684				
QY	541	GAGTTGACGCTCTCGCGCGCAACGGCGGACGGGTACAGAGGCTGGAGATCTCTGGAG	600				
DB	685	GAGTTGACGCTCTCGCGCGCAACGGCGGACGGGTACAGAGGCTGGAGATCTCTGGAG	744				
QY	601	GGCGCACCGAGTGTGTCTGAGCAACCTCGGGGGCGGACGCGCTACACCTTCGCCGTC	660				
DB	745	GGCGCACCGAGTGTGTCTGAGCAACCTCGGGGGCGGACGCGCTACACCTTCGCCGTC	804				
QY	661	CGCGCGGTATGGCTGAGCCGAGCTTCGGGGGCTTCTGGAGCGGCTGGTTCGGAGCCTGTG	720				
DB	805	CGCGCGGTATGGCTGAGCCGAGCTTCGGGGGCTTCTGGAGCGGCTGGTTCGGAGCCTGTG	864				
QY	721	TGCGTGTGACGCTTAGCGACCTTGACCCCTCATCTGACGCTTCTCCCTCATCTCCTGTG	780				
DB	865	TGCGTGTGACGCTTAGCGACCTTGACCCCTCATCTGACGCTTCTCCCTCATCTCCTGTG	924				
QY	781	GTCTATCTGTGTGTGACCGTCTGCGCTGTCTTCCACCGCGGGGCTCTGAAGCAG	840				
DB	925	GTCTATCTGTGTGTGCTGACCGTCTGCGCTGTCTTCCACCGCGGGGCTCTGAAGCAG	984				
QY	841	AAGATCTGGCTTGGCATCCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTTCACACCCAC	900				
DB	985	AAGATCTGGCTTGGCATCCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTTCACACCCAC	1044				
QY	901	AAGGGTAACCTTCAGCTGTGGCTGTACAGAAATGATGGCTGCTGTGTGGAGGCCCTTGC	960				
DB	1045	AAGGGTAACCTTCAGCTGTGGCTGTACAGAAATGATGGCTGCTGTGTGGAGGCCCTTGC	1104				
QY	961	ACCCCTTTCACGAGGAGCCACCTGCTTCCCTGGAAAGTCTCTCAGAGCGCTGTGGGGG	1020				
DB	1105	ACCCCTTTCACGAGGAGCCACCTGCTTCTCTGGAAAGTCTCTCAGAGCGCTGTGGGGG	1164				

```

Qy 1021 ACGATCAGCAGTGGAGCCGGGACAGATGATAGGGCCCCCTGCTGGAGCCAGTGGGC 1080
Db |||||
Qy 1165 ACGATCAGCAGTGGAGCCGGGACAGATGATAGGGCCCCCTGCTGGAGCCAGTGGGC 1224
Db |||||
Qy 1081 AGTGAGCATGCCAGGATACCTATCTGGTGTGGACAAATGGTTGTCGCCCGGAACCCG 1140
Db |||||
Qy 1225 AGTGAGCATGCCAGGATACCTATCTGGTGTGGACAAATGGTTGTCGCCCGGAACCCG 1284
Db |||||
Qy 1141 CCCAGTGAGGACCTCCAGGGCCTGGTGGCAGTGTGGACATAGTGGCCATGGATGAAGC 1200
Db |||||
Qy 1285 CCCAGTGAGGACCTCCAGGG----- 1305
Db |||||
Qy 1201 TCAGAGCATCCTCTCTGCTCATCTGCTTTGGCTCGAAGCCAGCCAGGGAGCCTCT 1260
Db ----- 1305
Qy 1261 GCTGCCAGCTTTGAGTACACTATCCTGGACCCAGCTCCAGCTCTTGGCTCCATGGACA 1320
Db |||||
Qy 1306 -----CCATGGGCA 1314
Db -----
Qy 1321 CTGTGCCCTGAGCTGCCCTACCCACCCACCTAAAGTACCTTGTGGTATCT 1380
Db |||||
Qy 1315 CTGTGCCCTGAGCTGCCCTACCCACCCACCTAAAGTACCTTGTGGTATCT 1374
Db |||||
Qy 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGCTTA 1440
Db |||||
Qy 1375 GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGCTTA 1434
Db |||||
Qy 1441 TCGGATGGCCCCCTACTCCAAACCTTATGAGAACAGCCCTTATCCAGCCCTGAGCCTCTG 1500
Db |||||
Qy 1435 TCGGATGGCCCCCTACTCCAGCCCTTATGAGAACAGCCCTTATCCAGCCCTGAGCCTCTG 1494
Db |||||
Qy 1501 CCCCCCAGCTATGTGGCTTGTCTTAG 1527
Db |||||
Qy 1495 CCCCCCAGCTATGTGGCTTGTCTTAG 1521

```

Search completed: April 20, 2005, 21:30:47
Job time : 856.5 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 19:31:21 ; Search time 6669 Seconds
(without alignments)

11094.784 Million cell updates/sec

Title: US-09-016-159D-4

Perfect score: 1527

Sequence: 1 atggaccacccggggcgctc.....gctatgtgcttctcttag 1527

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1527	100.0	1624	9 HUMERPR	M34986 Human eryth
2	1527	100.0	1818	6 AX658275	AX658275 Sequence
3	1527	100.0	1818	9 HUMERYTH	M50459 Human eryth
4	1525.4	99.9	1527	6 AR062465	AR062465 Sequence
5	1517.4	99.4	1848	6 CQ724319	CQ724319 Sequence
6	1312.2	85.9	1317	6 AR202361	AR202361 Sequence
7	1312.2	85.9	1317	6 AR202362	AR202362 Sequence
8	1307.4	85.6	1317	6 AR202362	AR202362 Sequence
9	1307.4	85.6	1317	6 AR202362	AR202362 Sequence
10	1276.6	83.6	2154	9 BC019092	BC019092 Homo sapi
11	1276.6	83.6	5761	9 AK074082	AK074082 Homo sapi
12	1155.6	75.7	1843	4 AF274305	AF274305 Sus scro
13	1080.6	70.8	1741	10 MUSERPR	J04843 Mouse eryth
14	1078.6	70.6	1695	10 BC003953	BC003953 Mus muscu
15	1075.8	70.5	1769	10 BC046282	BC046282 Mus muscu
16	1074.2	70.3	1859	10 S59388	S59388 Mus sp. ery
17	1058.2	69.3	1524	10 RATEPOR	D13566 Rattus norv
18	877.2	57.4	1164	4 AY029232	AY029232 Ovis arie
19	810.4	53.1	5683	6 AX554424	AX554424 Sequence

20	810.4	53.1	6122	6	AX554413	A
21	810.4	53.1	6122	6	AX554422	AX
22	810.4	53.1	6122	6	AX642149	AX6
23	810.4	53.1	6125	6	AX554420	AX5
24	810	53.0	4883	6	AX600116	AX50L
25	774.2	50.7	1257	4	AY029231	AY029
26	750.4	49.1	4990	6	AR119365	AR1193
27	750	49.1	750	6	AR031378	AR03137
28	750	49.1	750	6	BD134377	BD134377
29	750	49.1	750	6	AR428937	AR428937
30	750	49.1	750	6	BD009746	BD009746
31	748.2	48.0	6256	6	BD168201	BD168201
32	741	48.5	5565	6	BD168199	BD168199
33	702	46.0	1666	6	BD090961	BD090961
34	615.4	40.3	8647	9	S45332	S45332
35	615.4	40.3	119638	9	AC024575	AC024575
36	615.4	40.3	156897	2	AC018782	AC018782
37	615.4	40.3	217668	2	AC020561	AC020561
38	545	35.7	249366	2	AC151886	AC151886
39	524.4	34.3	686	4	BTU61399	BTU61399
40	524.4	34.3	686	4	BTU61399	BTU61399
41	517.2	33.9	951	10	D83509	D83509
42	514.4	33.7	824	6	E06723	E06723
43	478	31.3	236017	2	AC150437	AC150437
44	432.4	28.3	6561	10	MMERYPR	MMERYPR
45	432.4	28.3	216266	2	AC073797	AC073797

ALIGNMENTS

RESULT 1
HUMERPR
LOCUS HUMERPR 1624 bp mRNA linear PRI 08-NOV-1994
DEFINITION Human erythropoietin receptor mRNA, complete cds.
ACCESSION M34986.1
VERSION M34986.1 GI:182200
KEYWORDS erythropoietin receptor.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1624)
AUTHORS Ehrenman,K. and St John,T.
TITLE Sequence of the human erythropoietin gene
JOURNAL Unpublished (1990)
COMMENT Original source text: Human cell line OCI-M1, cDNA to mRNA.
Draft entry and computer-readable sequence for [1] kindly submitted by T.St.John, 07-JUN-1990.
Author address: T.St.John
Hutchinson Cancer Res Center
1124 Columbia St.
Seattle, WA
Email: TSTJOHN@FHCRVAX.

FEATURES

source
1..1624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="19p13.2"
1..1624
/genes="EPOR"
1..1527
/gene="EPOR"
/note="erythropoietin receptor precursor"

gene

/codon_start=1
/protein_id="AAA52401.1"
/db_xref="GI:182201"
/db_xref="GDB:G00-125-242"
/translation="MDHLGASLWPOVGSLLLAGAAWAPPPNLPDPKPKSKAALLAA
RPEELLCFTEERLEDLVCFWEERASAGVPGNGYSFYSQLEDFWKLRLHQPTARGA
RPFWCSLPATDSSFVPLERIVTAASGAPRYHRVTHINEVLLDAPVGLVARJADESG
HVYLRWLPPTPTMTSHIRYEDVSAGNGAGSVQVREIIEGRTECVLSNLRGRTRYTF

CDS

AVRAMEPFGFWSAMSEPVSLTLPDLPLILTLILVILVLLTLVLLSHRR
ALQKQIWPQIPSEFEGLFTTHKGNFOLWLYQNDGCLMWSPPCTPFTEDPPASLEVL
SERCKTMOAVBPQTDDEGLPLEPVGSEHAQDTYLVDKWLPRNPSPEDLPBGESV
DIVAMDEGSEASCSALAKSPSPGASASPEYILIDPSSOLLRLPWILCELPPTPP
HLKYLILVVSDDSGITDYSYSGDSQAQGLSDGPNFYNSLIIPAAEPLPLPSYVACS
"

sig_peptide

1..72

/gene="EPOR"

/note="erythropoietin receptor signal peptide"

73..1524

/gene="EPOR"

/product="erythropoietin receptor; G00-125-242"

ORIGIN

Query Match 100.0%; Score 1527; DB 9; Length 1624;
Best Local Similarity 100.0%; Pred. No. 1.5e-257;
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACCACTCGGGCGGTCCTCTGGCCCCAGGTGCGCTCCCTTTGTCTCTGCTCGCT 60
DB 1 ATGGACCACTCGGGCGGTCCTCTGGCCCCAGGTGCGCTCCCTTTGTCTCTGCTCGCT 60
QY 61 GGGGCGGCTGGGCGGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAAGCGCC 120
DB 61 GGGGCGGCTGGGCGGCCCCCGCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAAGCGCC 120
QY 121 TTGCTGGGCGGCCCCGGGCGGCAAGAGCTTCTGTCTTCAACGAGCGGTGGAGACTTG 180
DB 121 TTGCTGGGCGGCCCCGGGCGGCAAGAGCTTCTGTCTTCAACGAGCGGTGGAGACTTG 180
QY 181 GTGCTTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGGCGCGGCAATACAGCTTTCTCC 240
DB 181 GTGCTTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGGCGCGGCAATACAGCTTTCTCC 240
QY 241 TACCACTCGAGGATGACCATGGAAGCTGTGTGCTGCAACAGGCTCCACGCGCTGT 300
DB 241 TACCACTCGAGGATGACCATGGAAGCTGTGTGCTGCAACAGGCTCCACGCGCTGT 300
QY 301 GTGCGGCTGGCTTCTGGTGTTCGCTGCTACAGCGGACAGCTCGAGCTTCTGCGCCCTA 360
DB 301 GTGCGGCTGGCTTCTGGTGTTCGCTGCTACAGCGGACAGCTCGAGCTTCTGCGCCCTA 360
QY 361 GAGTTGCGGCTCACAGAGCTTCGGGCGCTCGCGATATACCGTGTCAATCAATCAAT 420
DB 361 GAGTTGCGGCTCACAGAGCTTCGGGCGCTCGCGATATACCGTGTCAATCAATCAAT 420
QY 421 GAAAGTAGTCTCTAGAGCGCCCGTGGGCTGGTGGCGCGGTGGCTGACGAGAGCGGC 480
DB 421 GAAAGTAGTCTCTAGAGCGCCCGTGGGCTGGTGGCGCGGTGGCTGACGAGAGCGGC 480
QY 481 CACGTAGTGTTCGCTGGCTCCCGCGGCTGAGACACCCATGACGTCTCACATCCGCTAC 540
DB 481 CACGTAGTGTTCGCTGGCTCCCGCGGCTGAGACACCCATGACGTCTCACATCCGCTAC 540
QY 541 GAGGTGAGAGCTTCGGCGGCGCAACCGGCGAGGAGCGGTACAGAGGTGGAGATCTGGAG 600
DB 541 GAGGTGAGAGCTTCGGCGGCGCAACCGGCGAGGAGCGGTACAGAGGTGGAGATCTGGAG 600
QY 601 GGGCGCACCGAGTGTGTGAGCAACCTGGGCGGCGGAGCGGTACACCTTCGCGCTC 660
DB 601 GGGCGCACCGAGTGTGTGAGCAACCTGGGCGGCGGAGCGGTACACCTTCGCGCTC 660
QY 661 CGCGCGCTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAGCGCTGTGCGAGCGCTGTG 720
DB 661 CGCGCGCTATGGCTGAGCCGAGCTTCGGCGGCTTCTGGAGCGCTGTGCGAGCGCTGTG 720
QY 721 TCGTGTGAGCGCTTACGAGCGCTGAGCCCGCTCATCTGTGAGCGCTTCCCTCATCTCGTG 780
DB 721 TCGTGTGAGCGCTTACGAGCGCTTACGAGCGCTTCCCTCATCTCGTG 780
QY 781 GTCATCTGGTGTCTGACGCTGTCTCGGCTGTCTCCACCGCGGCTCTGAGCAG 840
DB 781 GTCATCTGGTGTCTGACGCTGTCTCGGCTGTCTCCACCGCGGCTCTGAGCAG 840

QY 841 AAGATCTGGCTGGCATCCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTCCACCAACCAC 900
DB 841 AAGATCTGGCTGGCATCCCGAGCCCGAGAGCGAGTTTGAAGGCTCTTCCACCAACCAC 900
QY 901 AAGGTTAACTTCCAGCTGTGGCTGTACCAAGATGATGGCTGTGGTGGAGCCCTGCG 960
DB 901 AAGGTTAACTTCCAGCTGTGGCTGTACCAAGATGATGGCTGTGGTGGAGCCCTGCG 960
QY 961 ACCCCCTTACGGAGGAGCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGTGGGG 1020
DB 961 ACCCCCTTACGGAGGAGCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGTGGGG 1020
QY 1021 ACAGTCAGGAGGAGCGGAGACAGATGATGAGGCGCCCTCTGAGGCGAGTGGGC 1080
DB 1021 ACAGTCAGGAGGAGCGGAGACAGATGATGAGGCGCCCTCTGAGGCGAGTGGGC 1080
QY 1081 AGTCAGCATGCCAGGATACCTATCTGGTGTGGACAAATGTTGCTGCCCCGGAACCCG 1140
DB 1081 AGTCAGCATGCCAGGATACCTATCTGGTGTGGACAAATGTTGCTGCCCCGGAACCCG 1140
QY 1141 CCCAGTCAGGAGCTCCAGGCGCTTGGTGGCAGTGTGGCAGTGTGGCAGTGTGGATGAGGC 1200
DB 1141 CCCAGTCAGGAGCTCCAGGCGCTTGGTGGCAGTGTGGCAGTGTGGCAGTGTGGATGAGGC 1200
QY 1201 TCAGAACATCTCTCTGCTCATCTGCTTTGGCTTCGAAGCCCGAGGCGGAGCCCTCT 1260
DB 1201 TCAGAACATCTCTCTGCTCATCTGCTTTGGCTTCGAAGCCCGAGGCGGAGCCCTCT 1260
QY 1261 GCTGCCAGCTTTGAGTACATCTCTGGAGCCCGAGCTCCAGCTCTTGGCTCCATGACA 1320
DB 1261 GCTGCCAGCTTTGAGTACATCTCTGGAGCCCGAGCTCCAGCTCTTGGCTCCATGACA 1320
QY 1321 CTGTGCGCTCAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTTGTGCTATCT 1380
DB 1321 CTGTGCGCTCAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTTGTGCTATCT 1380
QY 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGAGCTCCAGGAGGCCAAGGGGCTTA 1440
DB 1381 GACTCTGGCATCTCAACTGACTACAGCTCAGGGAGCTCCAGGAGGCCAAGGGGCTTA 1440
QY 1441 TCGATGGCCCTTCTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCCTCTG 1500
DB 1441 TCGATGGCCCTTCTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCCTCTG 1500
QY 1501 CCCCCAGCTATGTGGCTTCTCTTAG 1527
DB 1501 CCCCCAGCTATGTGGCTTCTCTTAG 1527
RESULT 2
AX658275 1818 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 191 from Patent WO03000928.
DEFINITION AX658275
ACCESSION AX658275.1 GI:29160793
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
Petersen,M.W. and Elsnier,H.I.
TITLE Methods for identification of cancer cell surface molecules and
cancer specific promoters, and therapeutic uses thereof
JOURNAL Patent: WO 03000928-A 191 03-JAN-2003;
Odin Medical A/S (DK)
FEATURES
Location/Qualifiers
source 1..1818
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Qy	1021	ACGATCAGGACGAGTGAGCCGGGACAGATGATGAGGCCCCCTGCTGGAGACCAAGTGGGC	1081
Db	1126	ACGATCAGGACGAGTGAGCCGGGACAGATGATGAGGCCCCCTGCTGGAGACCAAGTGGGC	1185
Qy	1081	AGTGAGCATGCCAGGATACCTACTCTGGTGCTGGACAAATGGTTGCTGCCCGGAAACCCG	1140
Db	1186	AGTGAGCATGCCAGGATACCTACTCTGGTGCTGGACAAATGGTTGCTGCCCGGAAACCCG	1245
Qy	1141	CCGAGTGAGACCTCCAGGGCCCTGGTGCGAGTGTGGACANATGTGGCCATGGATGAAGGC	1200
Db	1246	CCGAGTGAGACCTCCAGGGCCCTGGTGCGAGTGTGGACANATGTGGCCATGGATGAAGGC	1305
Qy	1201	TCAGAGCATCCCTGCTGCTCATCTGCTTTGGCCCTCAAGCCAGCCAGAGGGAGGCTCT	1260
Db	1306	TCAGAGCATCCCTGCTGCTCATCTGCTTTGGCCCTCAAGCCAGCCAGAGGGAGGCTCT	1365
Qy	1261	GCTGCCAGCTTTGAGTACACTATCCTGGAGCCCCAGCTCCAGCTTTGCGTCCATGGACA	1320
Db	1366	GCTGCCAGCTTTGAGTACACTATCCTGGAGCCCCAGCTCCAGCTTTGCGTCCATGGACA	1425
Qy	1321	CTGTGCCCTGAGCTGCCCGCTACCCACCCACCTAAAGTACCTGTACCTTGTGGTATCT	1380
Db	1426	CTGTGCCCTGAGCTGCCCGCTACCCACCCACCTAAAGTACCTGTACCTTGTGGTATCT	1485
Qy	1381	GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCCAAGGGGCTTA	1440
Db	1486	GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCCAAGGGGCTTA	1545
Qy	1441	TCCGATGGCCCCCTACTCCAAACCCCTTATGAGAACAGCCTTATCCAGCCGCTGAGCCTCTG	1500
Db	1546	TCCGATGGCCCCCTACTCCAAACCCCTTATGAGAACAGCCTTATCCAGCCGCTGAGCCTCTG	1605
Qy	1501	CCCCCAGCTATGTGGCTTGCTCTTAG	1527
Db	1606	CCCCCAGCTATGTGGCTTGCTCTTAG	1632
RESULT 3			
HUMERYTH			
LOCUS	HUMERYTH 1818 bp mRNA linear PRI 08-NOV-1994		
DEFINITION	Human erythropoietin receptor mRNA, complete cds.		
ACCESSION	M60459		
VERSION	M60459.1 GI:182244		
KEYWORDS	erythropoietin receptor.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Jones, S.S., D'Andrea, A.D., Haines, L.L. and Wong, G.G.		
TITLE	Human erythropoietin receptor: cloning, expression, and biologic characterization		
JOURNAL	Blood 76 (1), 31-35 (1990)		
MEDLINE	90304340		
PUBMED	2163696		
COMMENT	Original source text: Human cDNA to mRNA.		
FEATURES	Location/Qualifiers		
source	1..1818		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/map="19p13.2"		
gene	1..1818		
	/gene="EPOR"		
CDS	106..1632		
	/gene="EPOR"		
	/codon_start=1		
	/product="erythropoietin receptor"		
	/protein_id="AA52403.1"		
	/db_xref="GI:182245"		
	/db_xref="GDB:G00-125-242"		
	/translation="MDHLGASLWPGVSLCLLAGAANAAPPNLDPDPKPFESKAALLAA		

ORIGIN

Query Match 99.9%; Score 1525.4; DB 6; Length 1527;
 Best Local Similarity 99.9%; Pred. No. 2.9e-257;
 Matches 1526; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGACCACTCGGGCGTCCCTCTGGCCCAAGTTCGGCTCCCTTTGTCTCTGCTCGCT 60
 Db 1 ATGGACCACTCGGGCGTCCCTCTGGCCCAAGTTCGGCTCCCTTTGTCTCTGCTCGCT 60

Qy 61 GGGGCGGCTGGGCGCCCGCCCTAACCTCCCGACCCCAAGTTCGAGAGCAAGCGGCC 120
 Db 61 GGGGCGGCTGGGCGCCCGCCCTAACCTCCCGACCCCAAGTTCGAGAGCAAGCGGCC 120

Qy 121 TTGCTGGCGGCGCCGGGCGCCGAAGAGCTTCTGTCTTCCAGGAGCGGTTCGAGACTTG 180
 Db 121 TTGCTGGCGGCGCCGGGCGCCGAAGAGCTTCTGTCTTCCAGGAGCGGTTCGAGACTTG 180

Qy 181 GTGTGTTTCTGGAGAAAGCGCGTGGGCTGCGGTGGGCGCCGGGCAACTACAGCTTCTCC 240
 Db 181 GTGTGTTTCTGGAGAAAGCGCGTGGGCTGCGGTGGGCGCCGGGCAACTACAGCTTCTCC 240

Qy 241 TACAGCTCGAGGATGAGCCATGGAAGCTGTGCGCTGCAACAGGCTCCACGCGCTCGT 300
 Db 241 TACAGCTCGAGGATGAGCCATGGAAGCTGTGCGCTGCAACAGGCTCCACGCGCTCGT 300

Qy 301 GGTGCGGTGCGCTTCTGTGTTTCTGCTGCTACAGCCGACAGCTCGAGCTTCTGCGCCCTA 360
 Db 301 GGTGCGGTGCGCTTCTGTGTTTCTGCTGCTACAGCCGACAGCTCGAGCTTCTGCGCCCTA 360

Qy 361 GAGTTGCGCGTCAACAGAGCTTCGGCGCTCCGGGATATACCGTGTCTATCCATCAAT 420
 Db 361 GAGTTGCGCGTCAACAGAGCTTCGGCGCTCCGGGATATACCGTGTCTATCCATCAAT 420

Qy 421 GAAAGTGTGCTCTAGAGCGCCCGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGG 480
 Db 421 GAAAGTGTGCTCTAGAGCGCCCGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGG 480

Qy 481 CACGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 481 CACGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

Qy 541 GAGTGGAGCTCTCGGCGGCAACGGCGAGGAGCGTACAGAGGTTGAGATCTTGAG 600
 Db 541 GAGTGGAGCTCTCGGCGGCAACGGCGAGGAGCGTACAGAGGTTGAGATCTTGAG 600

Qy 601 GCGCGACCGAGTGTGCTGAGCAACCTGCGGGCGGAGCGGCTACACCTTTCGCGCTC 660
 Db 601 GCGCGACCGAGTGTGCTGAGCAACCTGCGGGCGGAGCGGCTACACCTTTCGCGCTC 660

Qy 661 CGCGCGGTATGGCTGAGCGGCTTTCGGCGCTTTCGGAGCGCTGCTGCGAGCTGCTG 720
 Db 661 CTGCGCGGTATGGCTGAGCGGCTTTCGGCGCTTTCGGAGCGCTGCTGCGAGCGCTGCTG 720

Qy 721 TCGCTGCTGAGCGCTTACGAGCTTCGACCCCTCTATCTGAGCGCTTCTCCCTCATCTCGTG 780
 Db 721 TCGCTGCTGAGCGCTTACGAGCTTCGACCCCTCTATCTGAGCGCTTCTCCCTCATCTCGTG 780

Qy 781 GTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 GTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Qy 841 AAGATCTGGCTGGATCCCGAGCCCGAGAGCGAGTTCGAGCGCTTTCACCCACCCAC 900
 Db 841 AAGATCTGGCTGGATCCCGAGCCCGAGAGCGAGTTCGAGCGCTTTCACCCACCCAC 900

Qy 901 AAGGTAACTTCCAGCTGTGCTGTACCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 901 AAGGTAACTTCCAGCTGTGCTGTACCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

Qy 961 ACCCCCTTTCAGGAGGAGCCCACTGCTTCTTCTGGAAGTCTCTCAGAGCGCTGCTGCGGG 1020
 Db 961 ACCCCCTTTCAGGAGGAGCCCACTGCTTCTTCTGGAAGTCTCTCAGAGCGCTGCTGCGGG 1020

Qy 1021 ACGATGAGCAGTGGAGCGGGACAGATGATGAGGCGCCCTCTGCTGGAGCCAGTGGGC 1080
 Db 1021 ACGATGAGCAGTGGAGCGGGACAGATGATGAGGCGCCCTCTGCTGGAGCCAGTGGGC 1080

Qy 1081 AGTGAGATGCCAGGATACCTATCTGCTGCTGGACAAATGGTTGCTGCCCCGGAACCG 1140
 Db 1081 AGTGAGATGCCAGGATACCTATCTGCTGCTGGACAAATGGTTGCTGCCCCGGAACCG 1140

Qy 1141 CCCAGTGAGACCTCCAGGCGCTGCTGAGCATAGTGGCATAGTGGCATGGAAGGC 1200
 Db 1141 CCCAGTGAGACCTCCAGGCGCTGCTGAGCATAGTGGCATGGAAGGC 1200

Qy 1201 TCAGAACATCTCTCTGCTCATCTGCTTTGGCTTCGAAGCCAGCCAGAGGAGCGCTCT 1260
 Db 1201 TCAGAACATCTCTCTGCTCATCTGCTTTGGCTTCGAAGCCAGCCAGAGGAGCGCTCT 1260

Qy 1261 GCTGCCAGCTTTGAGTACATCTCTGAGCCAGCTCCAGCTCTTGGCTCCATGACA 1320
 Db 1261 GCTGCCAGCTTTGAGTACATCTCTGAGCCAGCTCCAGCTCTTGGCTCCATGACA 1320

Qy 1321 CTGTGCGCTGAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTTGTGATCT 1380
 Db 1321 CTGTGCGCTGAGCTGCGCCCTACCCACCCACCTAAAGTACCTGTACCTTGTGATCT 1380

Qy 1381 GACTCTGGCATCTCAACTGACTACGCTCAGGGAGCTCCAGGAGGCCAAGGGGGCTTA 1440
 Db 1381 GACTCTGGCATCTCAACTGACTACGCTCAGGGAGCTCCAGGAGGCCAAGGGGGCTTA 1440

Qy 1441 TCGATGGCGCCCTTCTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG 1500
 Db 1441 TCGATGGCGCCCTTCTCAACCTTATGAGAACAGCTTATCCAGCGCTGAGCCTCTG 1500

Qy 1501 CCCCCAGCTATGTGGCTTCTCTTAG 1527
 Db 1501 CCCCCAGCTATGTGGCTTCTCTTAG 1527

RESULT 5
 LOCUS CQ724319 1848 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 10253 from Patent WO02068579.
 ACCESSION CQ724319
 VERSION CQ724319.1 GI:42285176
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 KITS, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof
 Patent: WO 02068579-A 10253 06-SEP-2002;
 PE Corporation (NY) (US)

FEATURES
 source
 1..1848
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 99.4%; Score 1517.4; DB 6; Length 1848;
 Best Local Similarity 99.6%; Pred. No. 6.9e-256;
 Matches 1521; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGACCACTCGGGCGTCCCTCTGCGCCAGCTCGGCTCCCTTTGCTCTCTGCTCGCT 60
 Db 137 ATGACCACTCGGGCGTCCCTCTGCGCCAGCTCGGCTCCCTTTGCTCTCTGCTCGCT 196

Qy 61 GGGGCGGCTGGGCGCCCGCCCTAACCTCCCGAGCCCAAGTTCGAGAGCAAGCGGCC 120

197	GGGGCGCGCTGTGGCGGCCCCCGCGCTAACTCCCGGACCCCAAGTTTCAGAGAGCAAAAGCGGCG	255
121	TTGCTGGCGGCCCCGGGGGGCCGAAGAGCTTCTGTGCTTCCACGAGCGGTTGAGGACTTG	180
257	TTGCTGGCGGCCCCGGGGGGCCGAAGAGCTTCTGTGCTTCAACGAGGGTTGGAGACTTG	316
181	GTGTGTTTTCTGGAGGAAGCGCGAGAGCGTGTGGGTGGGCGCGGGCAAATTACAGCTTCTCC	240
317	GTGTGTTTTCTGGAGGAAGCGCGAGAGCGTGTGGGTGGGCGCGGGCAAATTACAGCTTCTCC	376
241	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCCGCTGCACACAGGCTCCGACGGCTCGT	300
377	TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTCCGCTGCACACAGGCTCCGACGGCTCGT	436
301	GGTGGGTGCGCTTCTGGTGTTCGTGTGCTCAGCGCGACACGTCGAGCTTCGTGCCCCCTA	360
437	GGTGGGTGCGCTTCTGGTGTTCGTGTGCTCAGCGCGACACGTCGAGCTTCGTGCCCCCTA	496
361	GAGTTGCGGCTCAGCAGAGCCTTCGGCGCTTCGGGATATACCGTGTATCCACATCAAT	420
497	GAGTTGCGGCTCAGCAGAGCCTTCGGCGCTTCGGGATATACCGTGTATCCACATCAAT	556
421	GAAGTAGTGCTCTTAGACGCCCCCGTGGGCTGTGTGGCGCGGTTGGCTTGACGAGAGCGCG	480
557	GAAGTAGTGCTCTTAGACGCCCCCGTGGGCTGTGTGGCGCGGTTGGCTTGACGAGAGCGCG	616
481	CACGTAGTGTTGCGTGTGCTCCGCGCGCTGAGACACCATGACGTCTCACATCCGCTAC	540
617	CACGTAGTGTTGCGTGTGCTCCGCGCGCTGAGACACCATGACGTCTCACATCCGCTAC	676
541	GAGTTGAGAGCTTCTGGCGGGCAACGGCGCAGGGAGCGTACAGAGGTGGAGTCTGTGAG	600
677	GAGTTGAGAGCTTCTGGCGGGCAACGGCGCAGGGAGCGTACAGAGGTGGAGTCTGTGAG	736
601	GGCCGCAACGAGTGTGCTCAGACAACTCGCGGGCGGACGCGTACACCTTCGCGCTC	660
737	GGCCGCAACGAGTGTGCTCAGACAACTCGCGGGCGGACGCGTACACCTTCGCGCTC	796
661	CGCGCGGTATGGCTGAGCGCGAGCTTCGGCGGCTTCTGGAGCGCTGTGTCGAGAGCTGTG	720
797	CGCGCGGTATGGCTGAGCGCGAGCTTCGGCGGCTTCTGGAGCGCTGTGTCGAGAGCTGTG	856
721	TCGCTGTGAGCGCTAGGCACTGGACCCCTCATCTCTGACGCTCTCCCTCATCTCTGTG	780
857	TCGCTGTGAGCGCTAGGCACTGGACCCCTCATCTCTGACGCTCTCCCTCATCTCTGTG	916
781	GTCACTCTGTGTGCTGACCGTCTCGCGTCTCTCCACCGCGGCTCTGAACGAG	840
917	GTCACTCTGTGTGCTGACCGTCTCGCGTCTCTCCACCGCGGCTCTGAAGGAG	976
841	AAGATCTGGCTTGGCATCCGAGCGCCAGAGCGAGTTTGAAGGCTCTTTCACCAACCCAC	900
977	AAGATCTGGCTTGGCATCCGAGCGCCAGAGCGAGTTTGAAGGCTCTTTCACCAACCCAC	1036
901	AAGGTTAACTTCCAGCTGTGCTGTACAGAAATGATGGCTGCCCTGTGGTGGAGCCCTTGC	960
1037	AAGGTTAACTTCCAGCTGTGCTGTACAGAAATGATGGCTGCCCTGTGGTGGAGCCCTTGC	1096
961	ACCCCTTTCAGGAGGACCACTGCTTCCCTGGAAGTCTCTTCAGAGCGCTGTGTGGGG	1020
1097	ACCCCTTTCAGGAGGACCACTGCTTCCCTGGAAGTCTCTTCAGAGCGCTGTGTGGGG	1156
1021	ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTTCTGGAGCCAGTGGCG	1080
1157	ACGATGAGGAGTGGAGCGGGGACAGATGATGAGGCGCCCTTCTGGAGCCAGTGGCG	1216
1081	AGTGAGCATGCCAGGATACCTATCTGGTGTGGAACAAATGGTTGTCTGCCCGGAACCCG	1140
1217	AGTGAGCATGCCAGGATACCTATCTGGTGTGGAACAAATGGTTGTCTGCCCGGAACCCG	1276
1141	CCAGTGAAGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC	1200
1277	CCAGTGAAGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGATGAAGGC	1336

Qy	1201	TCAGAGCATCCTCCTCGCTCATCTGCTTTGGCCTTCGAAGCCCGCAGAGGAGCGCTCT	1261
Db	1337	TCAGAAGCATCCTCCTCGCTCATCTGCTTTGGCCTTCGAAGCCCGCAGAGGAGCGCTCT	1396
Qy	1261	GCTGCCAGCTTTGAGTACATATCCTCTGGACCCCGAGCTCCCGAGCTCTTGGCTCCATGGACA	1320
Db	1397	GCTGCCAGCTTTGAGTACATATCTCTGGACCCCGAGCTCCCGAGCTCTTGGCTCCATGGACA	1456
Qy	1321	CTGTGCCCTGAGCTGCCCCCTACCCACCCACCTAAAGTACCTGTGTGGTATCT	1380
Db	1457	CTGTGCCCTGAGCTGCCCCCTACCCACCCACCTAAAGTACCTGTGTGGTATCT	1516
Qy	1381	GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCGAGGAGCCCAAGGGGGCTTA	1440
Db	1517	GACTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCGAGGAGCCCAAGGGGGCTTA	1576
Qy	1441	TCCGATGGCCCCCTACTCCAAACCTTATGAAACAGAGCTTATCCAGCGCTGAGCCTCTG	1500
Db	1577	TCCGATGGCCCCCTACTCCAAACCTTATGAAACAGAGCTTATCCAGCGCTGAGCCTCTG	1636
Qy	1501	CCCCCAGCTATGTGGCTTGCTCTTAG	1527
Db	1637	CCCCCAGCTATGTGGCTTGCTCTTAG	1663
RESULT 6			
AR202361			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Unclassified.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. .1317			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1314; Conservative			
0; Mismatches			
0; Gaps			
0;			
Qy	1	ATGAGCACCTCGGGGCGTCCCTCTGGCCCCCAGGTCGGCTCCCTTTGCTCTGCTCGCT	60
Db	1	ATGAGCACCTCGGGGCGTCCCTCTGGCCCCCAGGTCGGCTCCCTTTGCTCTGCTCGCT	60
Qy	61	GGGGCGCCTGGGCGCCCCCGCTAACTCCCGGACCCCAAGTTTCGAGACAAAGCGGC	120
Db	61	GGGGCGCCTGGGCGCCCCCGCTAACTCCCGGACCCCAAGTTTCGAGACAAAGCGGC	120
Qy	121	TTGCTGGGGCGGGGGCCGAGAGCTTCTGTCTTACCGAGGGTTGGAGACTTG	180
Db	121	TTGCTGGGGCGGGGGCCGAGAGCTTCTGTCTTACCGAGGGTTGGAGACTTG	180
Qy	181	GTGTGTTTCTGGAGGAAGCGGAGCGCTGGGGTGGGGCCGGGCAACTACAGCTTCTCC	240
Db	181	GTGTGTTTCTGGAGGAAGCGGAGCGCTGGGGTGGGGCCGGGCAACTACAGCTTCTCC	240
Qy	241	TACCAGCTCAGGATGAGCCATGGAAGTGTCTCGCTTSCACCAAGCTCCCAAGCTCGT	300
Db	241	TACCAGCTCAGGATGAGCCATGGAAGTGTCTCGCTTSCACCAAGCTCCCAAGCTCGT	300
Qy	301	GGTGGGTGGCTTCTGTGTTTGGCTTGCCTACAGCGGACAGTTCAGCTTGTGTCGCCCTA	360
Db	301	GGTGGGTAGCTTCTGTGTTTGGCTTGCCTACAGCGGACAGTTCAGCTTGTGTCGCCCTA	360

QY 361 GAGTTGCGGTCACAGAGCCTTCGGCGCTCCGGATATCACCGTGTATCCACATCAAT 420
 Db 361 GAGTTGCGGTCACAGAGCCTTCGGCGCTCCGGATATCACCGTGTATCCACATCAAT 420
 QY 421 GAAAGTAGTCTCTAGAGCGCCCGCTGGGCTGGGCGGTTGGCTGAGAGAGCGC 480
 Db 421 GAAAGTAGTCTCTAGAGCGCCCGCTGGGCTGGGCGGTTGGCTGAGAGAGCGC 480
 QY 481 CACGTAGTGTGCGTGGCTCCCGCGCTGAGACACCATGACGTCTCACATCGCTAC 540
 Db 481 CACGTAGTGTGCGTGGCTCCCGCGCTGAGACACCATGACGTCTCACATCGCTAC 540
 QY 541 GAGGTGAGCGTCTCGGCGGCAACCGCGAGAGCTTACAGAGGTTGAGATCTCGG 600
 Db 541 GAGGTGAGCGTCTCGGCGGCAACCGCGAGAGCTTACAGAGGTTGAGATCTCGG 600
 QY 601 GCGCGCAGAGTGTCTGAGCAACCTGGGCGGAGCGCTACACCTTCGCGTC 660
 Db 601 GCGCGCAGAGTGTCTGAGCAACCTGGGCGGAGCGCTACACCTTCGCGTC 660
 QY 661 GCGCGCAGTGTGAGCGAGCTTCGCGGCTTCTGAGCGCTGCTGAGAGCTGTG 720
 Db 661 GCGCGCAGTGTGAGCGAGCTTCGCGGCTTCTGAGCGCTGCTGAGAGCTGTG 720
 QY 721 TCGTGTGAGCGCTTACAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 780
 Db 721 TCGTGTGAGCGCTTACAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 780
 QY 781 GTCATCTGCTGCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 840
 Db 781 GTCATCTGCTGCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 840
 QY 841 AAGATCTGCTGCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 900
 Db 841 AAGATCTGCTGCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 900
 QY 901 AAGGTTAACTTCCAGCTGTGCTGTACAGAAATGATGCTGCTGCTGAGCGCT 960
 Db 901 AAGGTTAACTTCCAGCTGTGCTGTACAGAAATGATGCTGCTGCTGAGCGCT 960
 QY 961 ACCCTTTCAGGAGGAGCCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGG 1020
 Db 961 ACCCTTTCAGGAGGAGCCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGCTGGG 1020
 QY 1021 ACAGTACAGGAGTGGAGCGGAGCAGATGATGAGGCGCCCTGCTGAGCGCTGCTGGG 1080
 Db 1021 ACAGTACAGGAGTGGAGCGGAGCAGATGATGAGGCGCCCTGCTGAGCGCTGCTGGG 1080
 QY 1081 AGTGAGCATGCCAGGATACCTATCTGCTGTGAGCAAAATGTTGCTGCCCCGAAACCCG 1140
 Db 1081 AGTGAGCATGCCAGGATACCTATCTGCTGTGAGCAAAATGTTGCTGCCCCGAAACCCG 1140
 QY 1141 CCCAGTGGAGCCTCCAGGCGCTGCTGAGTGTGAGCATAGTGGCCATGATGAGGC 1200
 Db 1141 CCCAGTGGAGCCTCCAGGCGCTGCTGAGTGTGAGCATAGTGGCCATGATGAGGC 1200
 QY 1201 TCAGAGCATCTCTGCTGCTATCTGCTTGGCCTCGAAGCCAGCCAGAGGAGCTCT 1260
 Db 1201 TCAGAGCATCTCTGCTGCTATCTGCTTGGCCTCGAAGCCAGCCAGAGGAGCTCT 1260
 QY 1261 GCTGCCAGCTTGTAGTACATCTCTGAGACCCAGCTTCCAGCTTCTGCTGCTATG 1317
 Db 1261 GCTGCCAGCTTGTAGTACATCTCTGAGACCCAGCTTCCAGCTTCTGCTGCTATG 1317

RESULT 7
 AX008198
 LOCUS AX008198 1317 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 4 from Patent WO9967360.
 ACCESSION AX008198
 VERSION AX008198.1 GI:9995823
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Matthews,K.E., Bell,D. and Mueller,S.G.
 TITLE The efficient culture of stem cells for the production of
 hemoglobin
 JOURNAL Patent: WO 9967360-A 4 29-DEC-1999;
 HEMOSOL INC (CA)
 FEATURES Location/Qualifiers
 source 1..1317
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN 1315..1317
 Query Match 85.9%; Score 1312.2; DB 6; Length 1317;
 Best Local Similarity 99.8%; Pred. No. 6.3e-220;
 Matches 1314; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGACACACCTCGGGCGTCCCTCTGCCCCAGGTGGCTCCCTTTCTCTCTGCTGCT 60
 Db 1 ATGACACACCTCGGGCGTCCCTCTGCCCCAGGTGGCTCCCTTTCTCTCTGCTGCT 60
 QY 61 GGGGCGCCCTGGGCGCCCCCGCTTAACCTCCCGAGACCCCAAGTTTCGAGAGCAAGCGCC 120
 Db 61 GGGGCGCCCTGGGCGCCCCCGCTTAACCTCCCGAGACCCCAAGTTTCGAGAGCAAGCGCC 120
 QY 121 TTGCTGGCGCGCGGGGCGCGAGAGCTTCTGTGCTTACCGAGCGGTTGGAGACTTG 180
 Db 121 TTGCTGGCGCGCGGGGCGCGAGAGCTTCTGTGCTTACCGAGCGGTTGGAGACTTG 180
 QY 181 GTGTGTTTTCGGAGGAAGCGGCGCTGGGTGGGCGCGGCAACTACAGCTTCTCC 240
 Db 181 GTGTGTTTTCGGAGGAAGCGGCGCTGGGTGGGCGCGGCAACTACAGCTTCTCC 240
 QY 241 TACCAGCTCAGAGATGAGCCATGGAAGCTGTGCTGCTGACCAAGGCTCCACAGCGCTG 300
 Db 241 TACCAGCTCAGAGATGAGCCATGGAAGCTGTGCTGCTGACCAAGGCTCCACAGCGCTG 300
 QY 301 GGTGCGGTGCGCTTCTGCTGCTTACAGCCGACAGCTGAGCTTCTGCTGCGCCCTA 360
 Db 301 GGTGCGGTGCGCTTCTGCTGCTTACAGCCGACAGCTGAGCTTCTGCTGCGCCCTA 360
 QY 361 GAGTTGCGGCTCAGAGCAGCTCCGGCGCTCCGGATATCACCGTGTATCCACATCAAT 420
 Db 361 GAGTTGCGGCTCAGAGCAGCTCCGGCGCTCCGGATATCACCGTGTATCCACATCAAT 420
 QY 421 GAAAGTAGTCTCTAGACGCCCCCGTGGGCTGCTGGCGCGGTTGGCTGAGAGCGGC 480
 Db 421 GAAAGTAGTCTCTAGACGCCCCCGTGGGCTGCTGGCGCGGTTGGCTGAGAGCGGC 480
 QY 481 CAGGTAGTGTGGCTGGCTCCCGCGCTGAGACACCATGAGCTTCTACATCGGCTAC 540
 Db 481 CAGGTAGTGTGGCTGGCTCCCGCGCTGAGACACCATGAGCTTCTACATCGGCTAC 540
 QY 541 GAGGTGAGCTTCTGGCGGCAACCGCGCAGGAGCGTACAGAGGTTGAGATCTCGGAG 600
 Db 541 GAGGTGAGCTTCTGGCGGCAACCGCGCAGGAGCGTACAGAGGTTGAGATCTCGGAG 600
 QY 601 GGGCGCACCGAGTGTGTGCTGAGCAACCTGGGCGGAGCGCTACACCTTCGCGTC 660
 Db 601 GGGCGCACCGAGTGTGTGCTGAGCAACCTGGGCGGAGCGCTACACCTTCGCGTC 660
 QY 661 CCGCGCGTATGGCTGAGCGAGCTTCTGCGCGCTTCTGAGAGCGCTGCTCGGAGCCTGTG 720
 Db 661 CCGCGCGTATGGCTGAGCGAGCTTCTGCGCGCTTCTGAGAGCGCTGCTCGGAGCCTGTG 720
 QY 721 TCGCTGCTGAGCGCTAGCGACCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 780
 Db 721 TCGCTGCTGAGCGCTAGCGACCTGAGCGCTTCTGAGCGCTTCTGAGCGCTTCTGAGCGCT 780

Qy 1261 GCTGCCAGCTTTAGTACACTATCTCTGACCCAGCTCCAGCTCTTCCGTCCTCAGG 1317
Db 1261 GCTGCCAGCTTTAGTACACTATCTCTGACCCAGCTCCAGCTCTTCCGTCCTCAGG 1317

RESULT 9
LOCUS AX008201 1317 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 7 from Patent WO9967360.
ACCESSION AX008201
VERSION AX008201.1 GI:9995825
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Matthews,K.E., Bell,D. and Mueller,S.G.
The efficient culture of stem cells for the production of
hemoglobin
Patent: WO 9967360-A 7 29-DEC-1999;
JOURNAL HEMOSOL INC (CA)
FEATURES
source Location/Qualifiers
1..1317
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
terminator 1315..1317

ORIGIN
Query Match 85.6%; Score 1307.4; DB 6; Length 1317;
Best Local Similarity 99.5%; Pred. No. 4.4e-219;
Matches 1311; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ATGACACACCTCGGGCGTCCCTCTGCGCCAGCTCGGCTCCCTTTGTCCTCGCTCGCT 60
Db 1 ATGACACACCTCGGGCGTCCCTCTGCGCCAGCTCGGCTCCCTTTGTCCTCGCTCGCT 60

Qy 61 GGGGCCGCTGGGCGCCCGCCCTAACTCCCGGACCCCAAGTTTCGAGAGCAAGCGGCC 120
Db 61 GGGGCCGCTGGGCGCCCGCCCTAACTCCCGGACCCCAAGTTTCGAGAGCAAGCGGCC 120

Qy 121 TTGCTGGGGCGCCGGGGCCCGAAGAGCTTCTGTGCTTACCGAGCGGTTGGAGACTTG 180
Db 121 TTGCTGGGGCGCCGGGGCCCGAAGAGCTTCTGTGCTTACCGAGCGGTTGGAGACTTG 180

Qy 181 GTGTGTTTCTGGGAGGAGCGCGAGCGCTGGGCTGGGCGGCGCAACTACAGCTTCTCC 240
Db 181 GTGTGTTTCTGGGAGGAGCGCGAGCGCTGGGCTGGGCGGCGCAACTACAGCTTCTCC 240

Qy 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTGCTGCACCAAGCTCCACGGCTCGT 300
Db 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGTGCTGCTGCACCAAGCTCCACGGCTCGT 300

Qy 301 GGTGGGTGCGCTTCTGTGTGCTGCTGTACAGCGACAGCTCGAGCTTGTGTCGCCCTTA 360
Db 301 GGTGGGTGCGCTTCTGTGTGCTGCTGTACAGCGACAGCTCGAGCTTGTGTCGCCCTTA 360

Qy 361 GAGTTGGCGCTACAGAGCTTCCGGGCTCCGGATATCACGTGTCATCCACATCAAT 420
Db 361 GAGTTGGCGCTACAGAGCTTCCGGGCTCCGGATATCACGTGTCATCCACATCAAT 420

Qy 421 GAAGTAGTGTCTTACAGCGCCCGTGGGCTGTGTGCGGCTGTGCTGACGAGAGCGGC 480
Db 421 GAAGTAGTGTCTTACAGCGCCCGTGGGCTGTGTGCGGCTGTGCTGACGAGAGCGGC 480

Qy 481 CAGTAGTGTGTGCTGCTTCCGCCCGCTGAGACACCCATGAGCTTCTACATCCCGTAC 540
Db 481 CAGTAGTGTGTGCTGCTTCCGCCCGCTGAGACACCCATGAGCTTCTACATCCCGTAC 540

Qy 541 GAGGTGAGCTTCTGGCGCGCAAGCGCGAGGAGCTACAGAGGTTGGAGTCTTGGAG 600
Db 541 GAGGTGAGCTTCTGGCGCGCAAGCGCGAGGAGCTACAGAGGTTGGAGTCTTGGAG 600

Qy 601 GGCGCACCGAGTGTGTGCTGAGCAACCTGCGGGCCGAGCGCTACACCTTCCCGCTC 660
Db 601 GGCGCACCGAGTGTGTGCTGAGCAACCTGCGGGCCGAGCGCTACACCTTCCCGCTC 660

Qy 661 CGCGCGGTATGGCTGAGCGGCTTTCGGCGGCTTCTGGAGCGCTTGTGGAGCCTGTG 720
Db 661 CGCGCGGTATGGCTGAGCGGCTTTCGGCGGCTTCTGGAGCGCTTGTGGAGCCTGTG 720

Qy 721 TCGCTGCTGACGCTAGCGACCTTGGACCCCTCATCTGACGCTTCTCCCTCATCTCGT 780
Db 721 TCGCTGCTGACGCTAGCGACCTTGGACCCCTCATCTGACGCTTCTCCCTCATCTCGT 780

Qy 781 GTCATCTGCTGCTGACCTGCTGCGCTGCTCTCTCCACCGCGGCTTCTGAAGCAG 840
Db 781 GTCATCTGCTGCTGACCTGCTGCGCTGCTCTCTCCACCGCGGCTTCTGAAGCAG 840

Qy 841 AAGATCTGGCTGGCATCCGAGCCAGAGAGGAGTTTGAAGGCTTCTTCAACCCAC 900
Db 841 AAGATCTGGCTGGCATCCGAGCCAGAGAGGAGTTTGAAGGCTTCTTCAACCCAC 900

Qy 901 AAGGTTAACTTCCAGCTGTGGCTGTACCAATGATGGCTGCTGCTGAGGCTTCTG 960
Db 901 AAGGTTAACTTCCAGCTGTGGCTGTACCAATGATGGCTGCTGCTGAGGCTTCTG 960

Qy 961 ACCCTCTTACGAGGACCCACCTGCTTCTCTGGAAGTCTCTCAGAGCGCTTCTGG 1020
Db 961 ACCCTCTTACGAGGACCCACCTGCTTCTCTGGAAGTCTCTCAGAGCGCTTCTGG 1020

Qy 1021 ACATGAGGAGCTGGAGCGGGGACAGATGATGAGGCGCCCTCTGAGGAGCAGTGGC 1080
Db 1021 ACATGAGGAGCTGGAGCGGGGACAGATGATGAGGCGCCCTCTGAGGAGCAGTGGC 1080

Qy 1081 AGTGAGCATGCCAGGATACCTATCTGCTGCTGCAAAATGTTGCTGCGCCCGGAACCC 1140
Db 1081 AGTGAGCATGCCAGGATACCTATCTGCTGCTGCAAAATGTTGCTGCGCCCGGAACCC 1140

Qy 1141 CCAGTGAGGACCTCCAGGCGCTGTGGAGTGTGGACATGATGCGCATGGAAGGC 1200
Db 1141 CCAGTGAGGACCTCCAGGCGCTGTGGAGTGTGGACATGATGCGCATGGAAGGC 1200

Qy 1201 TCAGAGCATCTCTGCTCATCTGCTTGGCTCGAAGCCAGCCAGAGGAGCGCTCT 1260
Db 1201 TCAGAGCATCTCTGCTCATCTGCTTGGCTCGAAGCCAGCCAGAGGAGCGCTCT 1260

Qy 1261 GCTGCCAGCTTTGAGTACACTATCTCTGAGCCCGAGCTTCCAGCTTCTTGCCTCAT 1317
Db 1261 GCTGCCAGCTTTGAGTACACTATCTCTGAGCCCGAGCTTCCAGCTTCTTGCCTCAT 1317

RESULT 10
LOCUS BC019092 2154 bp mRNA linear PRI 02-JAN-2004
DEFINITION Homo sapiens erythropoietin receptor, mRNA (cDNA clone IMAGE:4894283), partial cds.
ACCESSION BC019092
VERSION BC019092.2 GI:40226074
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 2154)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,A.M., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stappleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Utschi,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2154)
Strausberg,R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:19387983.
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Nasello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Scantripop,S., Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.D., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: J Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557561.
FEATURES
source
1..2154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4894283"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH MGC 19"
/lab_host="pH10B-R"
/note="Vector: pOTB7"
ORIGIN
Query Match 83.6%; Score 1276.6; DB 9; Length 2154;
Best Local Similarity 99.7%; Pred. No. 9.9e-214;
Matches 1279; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 245 AGCTCGAGGATGAGGATGGAAGCTGTGTGCGCTGCAACAGGCTCCACGGCTCGGTGTG 304
DB 630 ATCGAGGGGATGAGGATGGAAGCTGTGTGCGCTGCAACAGGCTCCACGGCTCGGTGTG 689
QY 305 CGGTGGCTTCTGGTGTTCGCTGCTACAGCGGACAGCTCGAGCTTCGTGGCCCTTAGAGT 364
DB 690 CGGTGGCTTCTGGTGTTCGCTGCTACAGCGGACAGCTCGAGCTTCGTGGCCCTTAGAGT 749
QY 365 TCGGGCTCACAGCAGCTCCGGCGCTCCGGGATATACCGTGTATCCACATCAATGAAG 424
DB 750 TCGGGCTCACAGCAGCTCCGGCGCTCCGGGATATACCGTGTATCCACATCAATGAAG 809

QY 425 TAGTGCTCTAGACGCGCCCCCGTGGGGCTGTGGCGCGGTGGCTGACAGAGCGGCCACG 484
DB 810 TAGTGCTCTAGACGCGCCCCCGTGGGGCTGTGGCGCGGTGGCTGACAGAGCGGCCACG 869
QY 485 TAGTGTGGCTGTGGCTCCCGCGGCTGAGACACCCATGACCTCTCACATCCGCTACGAGG 544
DB 870 TAGTGTGGCTGTGGCTCCCGCGGCTGAGACACCCATGACCTCTCACATCCGCTACGAGG 929
QY 545 TGGACGCTCTCGCGCGGCAACCGGCGCAGGAGCGTACAGAGGCTGGAGATCTCTGGAGGGCC 604
DB 930 TGGACGCTCTCGCGCGGCAACCGGCGCAGGAGCGTACAGAGGCTGGAGATCTCTGGAGGGCC 989
QY 605 GCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGACGCGCTACACCTTCCGCGCTCCCGG 664
DB 990 GCACCGAGTGTGTGCTGAGCAACCTGCGGGGCGGACGCGCTACACCTTCCGCGCTCCCGG 1049
QY 665 CGCGTATGGCTGAGCGGAGCTTCCGGCGCTTCTCGAGCGCTGTGCGAGCGCTGTGTGCGC 724
DB 1050 CGCGTATGGCTGAGCGGAGCTTCCGGCGCTTCTCGAGCGCTGTGCGAGCGCTGTGTGCGC 1109
QY 725 TGCTGAGCGCTAGGAGCTGGACCCCTCATCTCTGAGCGCTCTCCCTCATCTCTCGTGTGCA 784
DB 1110 TGCTGAGCGCTAGGAGCTGGACCCCTCATCTCTGAGCGCTCTCCCTCATCTCTCGTGTGCA 1169
QY 785 TCCTGCTGTGCTGACCGTGTCTCGCGCTGTCTCCACCGCGCGGCTCTGAAGCAGAGA 844
DB 1170 TCCTGCTGTGCTGACCGTGTCTCGCGCTGTCTCCACCGCGCGGCTCTGAAGCAGAGA 1229
QY 1845 TCTGCGCTGGCATCCCGAGCGCAGAGCGAGTTTGAAGGCTCTTACACCAACCAACAGG 904
DB 1230 TCTGCGCTGGCATCCCGAGCGCAGAGCGAGTTTGAAGGCTCTTACACCAACCAACAGG 1289
QY 905 GTAACTTCCAGCTGTGGCTGTACAGAAATGATGCTGCTGCTGTGTGGAGCGCTCTGCACCC 964
DB 1290 GTAACTTCCAGCTGTGGCTGTACAGAAATGATGCTGCTGCTGTGTGGAGCGCTCTGCACCC 1349
QY 965 CTTTACGGAGGACCGACCTGTCTCCCTGGAAGTCTCTACAGCGCTGTGTGGGGAGCA 1024
DB 1350 CTTTACGGAGGACCGACCTGTCTCCCTGGAAGTCTCTACAGCGCTGTGTGGGGAGCA 1409
QY 1025 TGAGGAGCTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGAGCGAGCTGGGCGAGTG 1084
DB 1410 TGAGGAGCTGGAGCGGGGACAGATGATGAGGCGCCCTCTGCTGAGCGAGCTGGGCGAGTG 1469
QY 1085 AGCATGCCAGGATACCTATCTGTGTGTGGAATAATGGTGTGCTGCCCGGAACCGGCCCA 1144
DB 1470 AGCATGCCAGGATACCTATCTGTGTGTGGAATAATGGTGTGCTGCCCGGAACCGGCCCA 1529
QY 1145 GTGAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAG 1204
DB 1530 GTGAGGACCTCCAGGCGCTGGTGGCAGTGTGGACATAGTGGCCATGGATGAAGGCTCAG 1589
QY 1205 AAGCATCTCTGCTCATCTGCTTTGGCCTCGAAGCGCCAGCCAGAGGAGGAGCTCTGCTG 1264
DB 1590 AAGCATCTCTGCTCATCTGCTTTGGCCTCGAAGCGCCAGCCAGAGGAGGAGCTCTGCTG 1649
QY 1265 CCAGCTTTGATGATACCTATCTGTGAGCCCGCAGCTCTCCAGCTCTTGGTTCATGGAACATGT 1324
DB 1650 CCAGCTTTGATGATACCTATCTGTGAGCCCGCAGCTCTCCAGCTCTTGGTTCATGGAACATGT 1709
QY 1325 GCCCTGAGCTGCCCTTACCCCAACCCCTAAAGTACCTGTACCTTTGTGGTATCTGACT 1384
DB 1710 GCCCTGAGCTGCCCTTACCCCAACCCCTAAAGTACCTGTACCTTTGTGGTATCTGACT 1769
QY 1385 CTGCACTCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCGCCAAAGGGGGCTTATCCG 1444
DB 1770 CTGCACTCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCGCCAAAGGGGGCTTATCCG 1829
QY 1445 ATGCGCCCTACTCCAAACCCCTTATGAGAAAGCGCTTTATCCAGCGCGCTGAGCGCTCTGCCCC 1504
DB 1830 ATGCGCCCTACTCCAAACCCCTTATGAGAAAGCGCTTTATCCAGCGCGCTGAGCGCTCTGCCCC 1889
QY 1505 CCAGCTATGGGCTGTGCTCTTAG 1527

```
|||||
Db      1890 CCAGCTATGCGCTGCTTAG 1912

RESULT 11
AK074082
LOCUS      Homo sapiens mRNA for FLJ00153 protein.
DEFINITION
ACCESSION AK074082
VERSION    AK074082.1 GI:18676511
KEYWORDS   fls (full insert sequence).
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
TITLE      The nucleotide sequence of a long cDNA clone isolated from human
           spleen
JOURNAL    Published Only in Database (2002)
REFERENCE  2 (bases 1 to 5761)
AUTHORS    Jikuya,H., Takano,J., Nomura,N., Kikuno,R., Nagase,T. and Ohara,O.
TITLE      Direct Submission
JOURNAL    Submitted (21-JAN-2002) Takahiro Nagase, Kazusa DNA Research
           Institute, Department of Human Gene Research; 1532-3, Yana,
           Kisarazu, Chiba 292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp,
           URL:http://www.kazusa.or.jp/NEDO, Tel:81-438-52-3913,
           Fax:81-438-52-3914)
COMMENT    NEDO human cDNA sequencing project supported by Ministry of
           Economy, Trade and Industry of Japan; cDNA full insert and 5'- &
           3'-end one pass sequencing; Research Association for Biotechnology;
           cDNA library construction and clone selection; Kazusa DNA Research
           Institute
FEATURES             Location/Qualifiers
     source           1..5761
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FLJ00153"
                     /tissue_type="spleen"
                     /dev_stage="adult"
                     /note="vector:pBluescriptII SK plus"
     gene             1..5761
                     /gene="FLJ00153"
                     /evidence="not experimental"
                     /protein_id="BAB84908.1"
                     /db_xref="GI:18676512"
     CDS               1..5761
                     /gene="FLJ00153"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /evidence="not experimental"
                     /product="FLJ00153 protein"
                     /protein_id="BAB84908.1"
                     /translation="ARAGRAASGPPRVAQTSDDPSSEACABEEGLMPQQLDFES
                     VDEAQLTDLLELFKSVRYECGSVMSORDPFGAAGASPTVRATVAOFNTVTCV
                     LGSVLGAPLGAQRAQLEKWIIRIAQRELRNFSSRLAIRLALSQNPVYRLKRSWG
                     AVSRPLSTFKLSQIFSDENHLSREILFOEATEGSOEDNTPGSLSPKPPGPV
                     PVLGFTDLVLMETALPDMLEGLINFEKRWELILARIQOLQRCSQVTLSPHP
                     ILAHLAQNLITQSQRLSRVIEPPAASCSPSPRIRRLISLTVELSKLAREKSSP
                     SGSPDSSPSPTSSVSPSPSPSRDRAPGSPASPQSPQSPSTKLFLSLDLSPSPRF
                     ALPLGSPRIPLPQDSSEACAVIRVSDNDHGNLYRSLTSDKAPSVVRRALQKHNV
                     POPWACDYQLFQVLPGRGWEMLISAISVGTFFGFMQCARFYR"

ORIGIN
Query Match      83.6%; Score 1276.6; DB 9; Length 5761;
Best Local Similarity 99.7%; Pred. No. 8.2e-214;
Matches 1279; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      245 AGCTCGAGGATGAGCCATGGAGCTGTGCTGCGCTGCACAGGCTCCACAGGCTCGTGCTG 304
Db      3702 ATCGAGGGGATGAGCCATGGAGCTGTGCTGCGCTGCACAGGCTCCACAGGCTCGTGCTG 3761
Qy      305 CGGTGCGCTTCTGCTGTTGCTGCTTACAGCGGACACGCTCGAGCTTCGTGCCCCCTAGAGT 364
|||||
```



```

QY 1445 ATGGGCCCCCTACTCCAAACCCCTTATGAAACAGAGCTTATCCAGCCGCTGAGCCTCTGCCCC 1504
Db 4902 ATGGGCCCCCTACTCCAAACCCCTTATGAAACAGAGCTTATCCAGCCGCTGAGCCTCTGCCCC 4961

QY 1505 CCAGCTATGTGGCTTGTCTCTTAG 1527
Db 4962 CCAGCTATGTGGCTTGTCTCTTAG 4984

RESULT 12
AF274305
LOCUS
DEFINITION Sus scrofa erythropoietin receptor mRNA, complete cds.
ACCESSION AF274305
VERSION AF274305.1 GI:85722251
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 1843)
Pearson, P.L., Smith, T.P.L., Sonstegard, T.S., Klemcke, H.G.,
Christenson, R.K. and Vallet, J.L.
Porcine erythropoietin receptor: molecular cloning and expression
in embryonic and fetal liver
Domest. Anim. Endocrinol. 19 (1), 25-38 (2000)
20419811
10962196
2 (bases 1 to 1843)
Pearson, P.L., Smith, T.P.L., Sonstegard, T.S., Klemcke, H.G.,
Christenson, R.K. and Vallet, J.L.
Direct Submission
Submitted (01-JUN-2000) Reproduction Unit, Roman L. Hruska, U.S.
Meat Animal Research Center, P.O. Box 166, Clay Center, NE 68933,
USA

FEATURES
source
Location/Qualifiers
1..1843
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="liver"
/dev_stage="fetus"
160..1689
/note="cytokine receptor"
/codon_start=1
/product="erythropoietin receptor"
/db_xref="GI:85722252"
/protein_id="AAF77065.1"
/translation="MYHFGATLWPGVSLCLLAGATWAPSPNPDARFESKAALLAA
RGPEELICTERLEDVCFWEAGSAGVGPEDYSFYOEGEPKPKCHLHGPTARGS
VRFWCSLPDADTSFVPLELRYVTEVSSGAPRYHRIIHNEVLLDPPAGLLARAES
GHVLRWLPFGAPNASLIRVEVNIISTENAGGVQVREILDGRTECVLSNLRGTRVT
FMVRAARMAPSPFGFWMSAWSEPSALITDSDPLILTLULVLLILLALLSHR
RTLKQKIMPGIPSPGEPEGLFTTHKGNFQWLQYTDGCLWMSPTCPAEDPPAPLEV
LSERCWGVTOAVEAADDEGSLLEPVGSEHARDYLVLDKWLPPRRPASDELPOGGD
LDMAMDASASFCSSALALPKPEGASAAAFVETLDPSQLLRPALPAELPPTP
PHLKLYLVVSDGISDYSGGSQETQGSSSGSPYNSPYENSLVPAPEPSPNYVIC
S"

CDS
Query Match 75.7%; Score 1155.6; DB 4; Length 1843;
Best Local Similarity 85.2%; Pred. No. 1.6e-192;
Matches 1303; Conservative 0; Mismatches 224; Indels 3; Gaps 1;

QY 1 ATGACACACCTCGGGGCGTCCCTCTGCGCCCGCAGGTCGCGTCTCTGCTCGCTCGCT 60
Db 160 ATGATATCATTTTCGGGCAACCTCTGCGCAGAGTTGGCTCCCTCTGCTCTGCTAGCT 219

QY 61 GGGGCGGCTGGGCGCCCGCCCTAACCTCCCGGACCCCAAGTTCGAGAGCAAAAGCGGCC 120
Db 220 GGGGCGACCTGGGCTCTTACCCCAACTCCCGCCGACCAAGTTTGAAGCAAAAGCGGCC 279

```

```

QY 121 TTGCTGGCGGCCCGGGGGCCGAAAGAGCTTCTGTGCTTACCAGCGGTTGGAGACTTG 180
Db 280 CTGTTGGCAGCCCCGGGCGCTGAAGAGCTTCTGTGCTTACCAGCGGTTGGAGACTTG 339

QY 181 GTGTGTTTCTGGGAGGAAGCGGCGAGCGCTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
Db 340 GTGTGCTTCTGGGAGGAAGCGGGAAGCGCTGGAGTAGGCCCAAGAACTACAGCTTCTCC 399

QY 241 TACCAGCTCGAGGATGAGCCATGGAAGCTGTGCGCTGCAACAGGCTCCACAGGCTCGT 300
Db 400 TACCAGCTTGAGGGTGAGCCGCTGGAAGCATGCCACCTGCATCAAGGGCCCAAGGCCCGC 459

QY 301 GGTGCGGTGCGCTTCTGGTGTTCGCTGCTACAGCCGACAGCTCGAGCTTCTGTCGCCCTA 360
Db 460 GGTGCGGTGCGTTCGCTGCTGCTGCTACAGCCGACAGCTCGAGCTTCTGTCGCCCTA 519

QY 361 GAGTTGCGGCTCACAGCAG---CCTCGGGGCTCCGGGATATACCGTGTCTATCCACATC 417
Db 520 GAGCTGGCGGTCACTGAGGTTTCTCGGGGCTCCACGCTACCAACCGCTATCATCCACATC 579

QY 418 AATGAAGTAGTGCTCTAGACGCCGCCGCTGGGGCTGGTGGCGCGTGGCTGAGCAGAGC 477
Db 580 AACGAAGTGTGCTGTAGACCTTCCGCTGGGCTGTGGCAGCGCGGGCCGAAGAGAGT 639

QY 478 GGCCACGTAGTGTTCGCTGCTGCTGCTCCGCCGCTGAGACACCATGAGCTTTCACATCCG 537
Db 640 GGCCACGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 699

QY 538 TAGAGGTGAGAGCTTCTGGCGGGCAACGGCGCAGGAGCGGTACAGAGGTTGAGATCCGT 597
Db 700 TATGAGGTGAACATCTCGACAGAGAACCGCGCAGGGGCGGTGACAGAGGTTGGAATCTC 759

QY 598 GAGGGCCGACCCAGTAGTGCTGCTGAGCAACTGCGGGGCGCGGACCGCTACACCTTCGCC 657
Db 760 GACGGCCGCACTGAGTGTGCTGCTGAGCAACTGCGGGGTGGAACGCGCTACACCTTCATG 819

QY 658 GTCCGCGCGGTATGGCTGAGCGCGAGCTTCGGCGGCTTCTGGAGCGGCTGTGCGAGGCT 717
Db 820 GTACGCGCGGAAATGGCAGAGCGGAGCTTCGGTGGCTTCTGGAGCGGCTGTGTCGAGGCT 879

QY 718 GTGTGCTGTGCTGAGCGCTAGCGACCTCGACCCCGCTCATCTGAGCGCTCTCCCTCATCCTC 777
Db 880 GCGTGTGCTGTGACAGCTAGTGACTTGGACCCCGCTCATCTGACCCCTCTCCCTCATCCTC 939

QY 778 GTGTGCTATCTGTGTGCTGCTGACCGTCTCGCGCTGCTCTCCCAACCGCGGCTCTGAAG 837
Db 940 GTGTGCTATCTGTGTGCTGCTGCGCGTCTCGCGCTGCTCTCCCAACCGCTGAGGCT 999

QY 838 CAGAAGATCTGGCTTGGCTCCGAGCGCCAGAGAGCGAGTTTGAAGGCTCTTCAACACCC 897
Db 1000 CAGAAGATCTGGCTTGGTATCCCAAGCCCTGAGGGCGAGTTTGAAGGCTCTCTTCAACACCC 1059

QY 898 CACAAGGTAACCTTCCAGCTGTGCTGTACAGAACTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
Db 1060 CACAAGGTAATTTTCCAGCTGTGCTGTACCAAACTGATGATGATGATGATGATGATGATGATGAT 1119

QY 958 TGCAACCCCTTTCAGCGAGGACCCACCTGCTTCCCTGGAAGTCTCTCAGAGCGCTGTGCTG 1017
Db 1120 TGCACCCCTTTTCAGAGAGGACCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179

QY 1018 GGGACGATGAGGAGCTGGAGCGCGGGAACAGATGATGAGGCGCCCTCTGCTGAGGAGCAGTG 1077
Db 1180 GGGGTGACACAAGCAGTGAAGAACAGCGCGCAGATGATGAGGGGTCCTCTGCTGAGGAGCGGTG 1239

QY 1078 GGCAGTGAGCATGCCAGGATACCTATCTGTGCTGACAAATGTTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Db 1240 GGCAGTGAACATGCCCGAGACACCTACTGCTGTGCAAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1299

QY 1138 CCGGCCAGTGAAGACCTTCCAGGCGCTTGGTGGCAGTGTGGAATGTGGAATGTGGAATGTGGAATGTGGA 1197
Db 1300 CCGGCCAGTGAAGACCTTCCAGGCGCTTGGTGGTGTGATTTGGACATGGCGGCCATGGATGA 1359

QY 1198 GGCTCAGAAGCATCTCTCTGCTCATCTGCTTTGGCTTGGAGCCAGCCAGCCAGCCAGCCAGCCAGCC 1257

```


Db 1360 GCCTCAGAAGCATCTTCTGCTCATCTGCTTGGCCCTGAAAGCCTGGGCCAGAGGGGCC 1419
 Qy 1258 TCTGCTGCACGCTTGTAGTACATCTCTGACACCCAGCTCCAGCTCTTCTGCGTCCATGG 1317
 Db 1420 TCAGCTGCCAGCTTGTAGTATCTCTCTGATCTCCAGCTCAGAGCTCTTGGGCCCAAGA 1479
 Qy 1318 ACATGTGCTGAGCTGCCCCCTACCCACCCACCCCTAAAGTACCTGTACCTTGTGGTA 1377
 Db 1480 GCCCTGCCGCTGAGCTACCCCTACCCACCCACCCCTGAAGTACTTGTACTTGTGGT 1539
 Qy 1378 TCTGACTCTGCATCTCAACTGACTACAGCTCAGGAGCTCCAGGGAGCCAGGGGCC 1437
 Db 1540 TCTGACTCTGTCATCTCAACTGACTACAGCTCAGGGGCTCCAGGAAACCCAGGGGGC 1599
 Qy 1438 TTATCCGATGCCCTCTACTCCAAACCTTATGAGAACAGCCCTTATCCAGCGCTGAGCCCT 1497
 Db 1600 TCATCCAGTGGCCCTACTCCAAACCTTATGAGAACAGCCCTTATCCAGCGCTGAGCCCT 1659
 Qy 1498 CTGCCCCCAGCTATGTGGCTGTCTTAG 1527
 Db 1660 TCACCCCGAAGTATGTGACCTGCTCATAG 1689

RESULT 13

MUSERPR 1741 bp mRNA linear ROD 12-JUN-1993
 LOCUS Mouse erythropoietin receptor (EPO-R) mRNA, complete cds.
 DEFINITION J04843
 ACCESSION J04843.1 GI:193090
 VERSION erythropoietin receptor; hematopoietic growth factor receptor;
 KEYWORDS transmembrane protein.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1741)
 AUTHORS D'Andrea,A.D., Lodish,H.F. and Wong,G.G.
 TITLE Expression cloning of the murine erythropoietin receptor
 JOURNAL Cell 57 (2), 277-285 (1989)
 MEDLINE 89195238
 PUBMED 2539263
 COMMENT Original source text: Mouse erythroleukemia cell line MEL subclone 745, cDNA to mRNA, clone 190.
 Draft entry for [1] kindly provided by A.D. D'Andrea, 28-APR-1989.
 Location/Qualifiers
 1..1741
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 28..1551
 /notes="erythropoietin receptor precursor"
 /codon_start=1
 /protein_id="AA337571.1"
 /db_xref="GI:309219"
 /translation="MOKLRVPLRVGRLCLLAGAAWAPSPSPDPKFSKAAALAS
 RGSSEILLCFTDSFVLEWEEAASGMDNFYSQLEGESEKSGSLHQAPTVRSV
 RFWCSILPTADTSFVLEWEEAASGMDNFYSQLEGESEKSGSLHQAPTVRSV
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 LOOKIWPGPSPESEFEGFTTHKGNFQLMQLRDGLMWSPGSPFPDPPAHLEVL
 EPRWAVTQADGPDGEPGLLEPVGSEHAQDTYLVLDKLLPRTCPSENLSGPGSDV
 PTWDBASETSVCPDLASKPRPEGTSPSPFEVITLDPSOLLCPRALPELPPTPPH
 LKYLVLVPSDSGISTDYSSGSGQVHGDSGYPHYENSLVDPDSEPLHPGVACS"
 28..159
 /notes="erythropoietin receptor signal peptide"
 100..1548
 /product="erythropoietin receptor"
 /region="271 bp upstream of XhoI site."

FEATURES

source
 1..1741
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 28..1551
 /notes="erythropoietin receptor precursor"
 /codon_start=1
 /protein_id="AA337571.1"
 /db_xref="GI:309219"
 /translation="MOKLRVPLRVGRLCLLAGAAWAPSPSPDPKFSKAAALAS
 RGSSEILLCFTDSFVLEWEEAASGMDNFYSQLEGESEKSGSLHQAPTVRSV
 RFWCSILPTADTSFVLEWEEAASGMDNFYSQLEGESEKSGSLHQAPTVRSV
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 VLRWLPPEAPMTTHIRYEDVYSAGNRAGTORVEVLEGEVTECVLSNLRGTRYFA
 LOOKIWPGPSPESEFEGFTTHKGNFQLMQLRDGLMWSPGSPFPDPPAHLEVL
 EPRWAVTQADGPDGEPGLLEPVGSEHAQDTYLVLDKLLPRTCPSENLSGPGSDV
 PTWDBASETSVCPDLASKPRPEGTSPSPFEVITLDPSOLLCPRALPELPPTPPH
 LKYLVLVPSDSGISTDYSSGSGQVHGDSGYPHYENSLVDPDSEPLHPGVACS"
 28..159
 /notes="erythropoietin receptor signal peptide"
 100..1548
 /product="erythropoietin receptor"
 /region="271 bp upstream of XhoI site."

ORIGIN

271 bp upstream of XhoI site.
 Query Match 70.8%; Score 1080.6; DB 10; Length 1741;
 Best Local Similarity 82.2%; Pred. No. 2.2e-179;
 Matches 1255; Conservative 0; Mismatches 269; Indels 3; Gaps 1;

Qy 1 ATGGAACACCTCGGGCGTCCCTCTGGCCCCAGGTGCGCTCCCTTTGTCTCTCTCGCT 60
 Db 28 ATGGAACAACTCAGGGTGGCCCTCTGGCCCTCGGTAGGCCCTCTCTCTCTCTCTCT 87
 Qy 61 GGGGCGCTGGGGCGCCCGCCCTTAACCTCCCGGACCCCAAGTTCGAGAGCAAGCGGCC 120
 Db 88 GGGGCGAGCTGGGCACTCTTCAACCCAGCTCCCGGACCCCAAGTTCGAGAGCAAGCGGCC 147
 Qy 121 TTCTGCGCGCCCGGGGCGGAGAGCTTCTGTCTTACCGAGCGGTTGGAGACTTG 180
 Db 148 CTGCTGTCATCCCGGGCTCCGAAAGAACTTCTGTCTTACCCCAACGCTTGGAGACTTG 207
 Qy 181 GTGTGTTTTCGGAAGAAAGCGGAGCGTGGGGTGGGCCCGGGCAACTACAGCTTCTCC 240
 Db 208 GTGTGTTTTCGGAAGAAAGCGGAGCTCGGGATGGAC--TTCAACTACAGCTTCTCA 264
 Qy 241 TACCACTCAGAGTACAGCATGGAAGCTGTGTGGCTGACCAAGGCTCCACGGCTCGT 300
 Db 265 TACCACTCAGAGGTGAGTACGAAAGTCAATGTAGCTTGCACCAAGGCTCCACGGCTCCG 324
 Qy 301 GGTGCGGTGCGCTTCTGGTGTTCGCTTACAGCGACACAGCTCGAGCTTCTGTCGCCCTA 360
 Db 325 GGCTCGTGGTTCGTGGTTCACAGCGACACAGCGGACACATCGAGTTTGTGCGCGTG 384
 Qy 361 GAGTTGCGGTTCACAGCAGCCTCCGGCGTCCCGGATATCACCGTGTCTATCCACATCAAT 420
 Db 385 GAGTCTCAGGTGACGAGGCGTCCGGTTCCTCTCGCTATCACCGCATCATCCATATCAAT 444
 Qy 421 GAAGTAGTCTCTAGACGCCCGCTGGGGCTGGTGGCGGGTGGCTGAGGAGGCGG 480
 Db 445 GAAGTAGTCTCTGAGCGCCCGCGGGGCTGCTGGCGCGGCGGAGAGGCGGAGC 504
 Qy 481 CAGTAGTGTGGCTGGCTCCCGCGCTCGAGACACCCATGACCTCTACATCCCGCTAC 540
 Db 505 CAGTGTGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 564
 Qy 541 GAGGTGAGCTCTCGGCCCGGCAACCGCGGAGGAGCGTACAGAGGTGGAGATCTCTGGAG 600
 Db 565 GAAGTGGAGCTGCGGCGAGCAACCGCGGAGGAGGAGCAACAAAGGCTGGAGGCTCTGGAA 624
 Qy 601 GCGCGCACGAGTGTGTGTGAGCAACCTCGGGCGGCGGAGCGCTACACCTTCGCGCTG 660
 Db 625 GCGCGCACGAGTGTGTGTGAGCAACCTCGGGCGGCGGAGCGCTACACCTTCGCGCTG 684
 Qy 661 CGCGCGCTATGGCTGAGCGGAGCTTCTGGCGGCTTCTGGAGCGCTGGTGGAGCGCTGTG 720
 Db 685 CGAGCGGATGGCTGAGCGGAGCTTCTGGAGCTTCTGGAGCTGGTGGTGGTGGTGGT 744
 Qy 721 TCGCTCTGACGCTAGCGACCTGGACCCCTCATCTTCTGAGCGCTCTCTCTCTCTCTCT 780
 Db 745 TCACTACTGACGCTAGCGACCTGGACCCCTCTCATCTTGAACGCTGTCTCTCTCTCTCT 804
 Qy 781 GTCACTCTGGTGTGTGTGAGCGCTGCTCGCGCTGCTCTCCACCGCGGGCTCTGAGCAG 840
 Db 805 CTCACTCTGCTGTGTGTGAGCGCTTCTGGCGCTGCTCTCCACCGCGGGAGCTCTGAGCAG 864
 Qy 841 AAGATCTGCGCTGGCATCCGAGCGGAGCGAGTGTGAGAGCTCTTTCACCAACCCAC 900
 Db 865 AAGATCTGCGCTGGCATCCGAGCGGAGCGAGTGTGAGAGCTCTTTCACCAACCCAC 924
 Qy 901 AAGGTGAACTTCCAGCTGGCTGTGACCAAGATGATGGTGTCTGTGGTGGAGCGCTCTG 960
 Db 925 AAGGTGAACTTCCAGCTGGCTGTGACCAAGATGATGGTGTCTGTGGTGGAGCGCTCTG 984
 Qy 961 ACCCTCTCAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1020
 Db 985 AGCT 1044
 Qy 1021 ACATGTCAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1080
 Db 1045 GTGACTCAGGCTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1104

121 GCTGTCATCCCGGGGCTCCGAGAACTTCTGTGCTTCAACCAACGCTTGGAGACTTGGT 180
183 GTGTTTCTGGGAGGAGCGGAGCGCTGGGCTGGGCGCGGCAACTACAGCTTCTCCTA 242
181 GTGTTTCTGGGAGGAGCGGAGCTCCGGATGAC---TTCACTACAGCTTCTCAT 237
243 CCAGCTCAGAGATGAGCATGGAAGCTGTGTGCGCTGCAACAGGCTCCCAAGGCTGCG 302
238 CCAGCTCAGAGGTTAGTACGAAAGTCAATGTAGCTGCAACAGGCTCCCAAGGCTCG 297
303 TCGGCTGCGCTTCTGCTGCTCGCTACAGCGACACGTCGAGCTTCGTGCGCCCTAGA 362
298 CTCGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
363 GTTGGCGCTCAGCAGCGCTCCGGCGCTCCGCGATATCACCGTGTCAATCAATGA 422
358 GCTGAGGTGAGCGGAGCGCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
423 AGTAGTCTCTAGACGCGCGCTGGGCTGGTGGCGGTTGGTGTGACGAGCGGCA 482
418 AGTAGTCTCTGAGCGCGCGCGGCTGCTGGCGCGCGGCGGCAAGAGGCGGCA 477
483 GGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
478 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
543 GGTGAGCGCTTCTGCGCGCGCAACGCGGAGGAGGCTACAGAGGCTGAGATCTTGGAGG 602
538 AGTGAGCGTGTGCGGAGCAACCGGCGAGGAGGACAAAGGCTGAGGCTTCTGGAAG 597
603 CCGCACCGAGTGTGTGTGAGCAACCTGCGGGCGGAGCGGCTACACCTTTCGCGTCCG 662
598 CCGCACCTGAGTGTGTGTGAGCAACCTGCGGGCGGAGCGGCTACACCTTTCGCTGTC 657
663 CCGCGTATGCTGAGCGGAGCTTGGGGGCTTCTGAGGCGCTTCTGAGGCGCTGCTGTC 722
658 AGCGCGATGCGCGGAGCGGAGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 717
723 GCTGCTGAGCGCTGAGCGGAGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 782
718 ACTACTGAGCGCTGAGCGGAGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 777
783 CATCTCGTGTGTGAGCGGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 842
778 CATCTCGTGTGTGAGCGGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 837
843 GATCTGCGCTGAGCGGAGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 902
838 GATCTGCGCTGAGCGGAGCTTCTGAGGAGCTTCTGAGGCGCTTCTGAGGCGGCT 897
903 GGTAACTTCCAGCTGTGGCTGTACCAAGATGATGGCTGCTGCTGCTGCTGCTGCTGCT 962
898 GGTAACTTCCAGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
963 CCCCTTCCAGGAGCGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
958 CTCCTTCCCTGAGGATCCAGCTGCGGAGCTTCTGAGGAGCTTCTGAGGAGCTTCTGAG 1017
1023 GATGCGGAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1082
1018 GATGCGGAGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1077
1083 TGAGCATGCCAGGATACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
1078 TGAGCATGCCAGGATACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
1143 CAGTGAGGAGCTCCAGGCGCTGTGGCAGTGTGACATAGTGCCCATGATGAGGCTC 1202
1138 CAGTGAGGAGCTTCTCAGGCGCTGGGCGAGTGTGACCCCTGTGACTATGATGAGGCTC 1197
1203 AGAAGCATCTCTGCTATCTGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT 1262

Db 1198 AGAAACATCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
Oy 1263 TGCAGCTTTGAGTACACTATCTTGGACCCAGCTCCAGCTCTTGGTCTCATGACACT 1322
Db 1258 TTCCAGCTTTGAGTACACTATCTTGGACCCAGCTCTGAGCTCTGCTGCTGCTGCTGCT 1317
Oy 1323 GTGCGCTGAGCTGCGCTTACCCACCCACCCACCTAAAGTACTGTTGCTGCTGCTGCTGCT 1382
Db 1318 GCTTCCCGAGTACCTCCCACTCCACCTCACTTGAAGTACTATACCTTGTGCTGCTGCTGCT 1377
Oy 1383 CTCTGGCATCTCAACTGACTACAGCTCAGGGGACTCCAGGGAGCCCAAGGGGCTTATC 1442
Db 1378 TTCTGGCATCTCAACAGATTACAGTTCCGGGGGCTCTCAGGGAGTCCACGGGAGTCTATC 1437
Oy 1443 CGATGGCCCTACTTCCAAACCTTATGAGAACAGCTTATCCAGCGCTGAGCGCTCTGCTGCT 1502
Db 1438 TGATGGCCCTACTTCCACCCCTATGAGAACAGCTTGTGCCAGACTCAGAGCTCTGCTGCT 1497
Oy 1503 CCCAGCTATGTGGCTTGTCTCTTAG 1527
Db 1498 TCCCGGCTATGTGGCTGCTCTTAG 1522

RESULT 15
BC046282
LOCUS
DEFINITION Mus musculus erythropoietin receptor, mRNA (cdna clone MGC:54669 IMAGE:6438591), complete cds.
ACCESSION BC046282
VERSION BC046282.1 GI:28279360
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1769)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywinski,M.I., Skalka,U., Smalley,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1769)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgenev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 100 Row: i Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6753765.

FEATURES

source

1..1769
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="MGC:54669 IMAGE:6438591"
/tissue type="Mammary tumor. MMTV-LTR/INT3 model. 5 month
old mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1769
/gene="Epor"
/db_xref="LocusID:13857"
/db_xref="MGI:95408"
76..1599
/gene="Epor"
/codon_start=1
/product="Epor protein"
/protein_id="AAH46282.1"
/db_xref="GI:28279361"
/db_xref="LocusID:13857"
/db_xref="MGI:95408"
/translation="MDKLRLVPLWPRVGLCLLLLAGAAWAPSPPLDPKFSKALLAS
RGSEELICFTORLEDLFWEEAASSGDMFNYSYOLEGESRKSLHQAPTVRGVS
RFWCSLTADPTSPVLEQVTEASGSPRYHRIHINEVLLDAPAGLLARABEGSH
VWLRLPPGAPMTTHIRYEDVAGNAGRGATQREVELEGTECVLNLRGRTYTPA
VRAMAPSRFSGSAMSEPAASLITASDLPLILTLILVLSLITVLLALSHRRT
LOKIMWPIPSPESEFSGLPFTYHKNFQLMLLQDGLMWSPGSSFPEDPPAHLEVL
EPRWAVTQADPGADDDGPLEPVGSEHAQDYLVDKWLPLPTPCSENLSGPGGSD
PVTWDEASPTSSCPDLSAKPRPEGTSPSSFEYITLDPSSQLLCFRALPPELPTPH
LKYLYLVVSDSGISTDYSSGSGQVHGSDSPGYSHFYENSLVDPSEPLHPGVACS"
508..741
/gene="Epor"
/note="FN3; Region: Fibronectin type 3 domain"
/db_xref="CDD:smart00060"

gene

CDS

misc_feature

ORIGIN

Query Match 70.5%; Score 1075.8; DB 10; Length 1769;
Best Local Similarity 82.0%; Pred No. 1.5e-178;
Matches 1252; Conservative 0; Mismatches 272; Indels 3; Gaps 1;
Qy 1 ATGGACCACTCGGGGCGTCCCTCTGGCCCGCAGTTCGGCTCCCTTTGTCTCTCGCTCGCT 60
Db 76 ATGGACAACTCAGAGTGCCTCTGGCTCGGGTAGGCCCTCTGTCTCTACTTGTCT 135
Qy 61 GGGGCGGCTGGGGCGCCCGCCCTAACTCCCGGACCCCAAGTTGCGAGAGCAAGCGGCC 120
Db 136 GGGGCGGCTGGGCACTTCAACCGAGCTCCCGGACCCCAAGTTGAGAGCAAGCGGCC 195
Qy 121 TTGCTGGGGCGGGGCGGAGAGCTTCTGTGCTTCAACGAGCGGTTGAGAGCTTG 180
Db 196 CTGCTGGCATCCGGGGGCTCGAAGAACTTCTGTGCTTCAACCGAGCTTGGAGAGACTTG 255

Qy 181 GTGTGTTTCTGGAGGAAGCGCGAGCGCTGGGGTGGCCCGGCAACTACAGCTTCTTC 240
Db 256 GTGTGTTTCTGGAGGAAGCGCGAGCTCCGGATGGAC--TTCAACTACAGCTTCTCA 312
Qy 241 TACCAGCTCAGAGTAGCCATGGAAGCTGTGTCCCTGACACGAGGTCCTCCACGCGCTCGT 300
Db 313 TACCAGCTCAGAGTAGCTCACCAGAGTCAATGAGCTGTGACACGAGGTCCTCCACGCTCCGC 372
Qy 301 GGTGCGGTGGCGTCTCTGGTGTTCGCTGCTTACAGCCGACACGCTCGAGCTTCTGTCGCCCTA 360
Db 373 GGCTCCGTGGTGTCTGTGTGTCTCACTGCCAAAGCGGACACATCGAGTCTTGTGCGCGTG 432
Qy 361 GAGTTGCGCTCAGAGAGCTCTCGGCGCTCCGCGATATCACCGTGTCTCAATCAATCAAT 420
Db 433 GAGCTGAGCTGAGCGAGGCTCCGGTCTCTCGCTATCACCGCATCATCCATATCAAT 492
Qy 421 GAGTAGTGTCTTACAGCGCCCGCTGGGCTGGTGGCGGTTGGCTGACGAGAGCGGC 480
Db 493 GAGTAGTGTCTTGGACGCCCGCGGCGCTCTGCGCGCGCGGCGAGAGGCGGCGC 552
Qy 481 CAGCTAGTGTGGCTGGCTCCGCGCGCTTGAGACACCCATGACGTCTCATCTCCGCTAC 540
Db 553 CAGCTGCTGCTGGCTGCCACCTCTCTGGAGCACCTATGACCCACATCCGATAT 612
Qy 541 GAGTTGAGCTCTCGGCGCGCAACGGCGAGGAGCGGTACAGAGGTGGAGATCTCTGGAG 600
Db 613 GAAAGTGAGCTGTGCGGAGGCAACCGGCGAGGAGGAGGACACAAAGGTTGGAGTCTGAGAA 672
Qy 601 GCGCGCACGAGTGTGTGCTGACCACTCGCGGCGGAGCGGTACACTTCGCGCGTC 660
Db 673 GCGCGCACTGAGTGTGTCTGAGCAACCTCGCGGCGGAGCGGTACACTTCGCTGTT 732
Qy 661 CGCGCGGTATGGCTGAGCGGAGCTTTCGCGGCTTCTGGAGCGCTTGGTTCGAGCGCTGTG 720
Db 733 CGAGCGCATGCCGAGCGGAGCTTTCAGCGGATTTCTGGAGTGCCTGGTCTGAGCGCGC 792
Qy 721 TCCTGCTGAGCGCTAGCGACCTGGAGCCCTCATCTGAGCGCTCCCTCATCTCTGCTG 780
Db 793 TCACTAGTACCGCTAGCGACCTGGAGCCCTCTCATCTTGAAGCTGTCTCTCATCTCTGCTC 852
Qy 781 GTCATCTGGTGTGTGCTGACCGTGTCTCGCGCTCTCTCCACCGCGCGGCTCTGAAAGCAG 840
Db 853 CTATCTCTGTGTGTGCTGAGGGTCTTGGCCCTGCTGTCCACCGCGGACTCTGCGAGCAG 912
Qy 841 AAGATCTGGCTGGCATCCGAGCCAGAGCGAGGTTTGAAGCGCTCTTCAACACCCAC 900
Db 913 AAGATCTGGCTGGCATCCAGCCAGAGCGAGTTTGAAGGCTCTTCTTCAACACCCAC 972
Qy 901 AAGGTTAACTTCCAGCTGTGGCTGTACCAAGATGATGGCTGCCTGTGTTGGAGAGCCCTGC 960
Db 973 AAGGTTAACTTCCAGCTGTGGCTGTGCGAGCGGTGATGGTGTCTGTGTTGGAGAGCCCGGC 1032
Qy 961 ACCCTTTCAGGAGGACCCACCTGTCTTCCCTGGAAAGTCTCTCAGAGCGCTGTCTGGGG 1020
Db 1033 AGTCTCTTCCCTGAAGATCCACTGCTGCCACCTAGAGGTCTCTCAGAGCCACGCTGGGCA 1092
Qy 1021 ACGATCAGGAGCTGAGCGCGGGAAGATGATGAGGCGCCCTCTGCTGAGAGCAGTGGGC 1080
Db 1093 GTTACTCAGGCTGGGAGACCCAGGCGGAGATGATGAGGCGCCCTTACTGAGAGCGGTGGGC 1152
Qy 1081 AGTGAGCATGCCAGATACCTATCTGTGCTGGAAGAAATGGTGTGCTGCGCCCGGAAACCCG 1140
Db 1153 AGTGAGCATGCCAGGACACCTACTTGGTATTGGATTAAGTGGTGTGCTGCGCCCGGACCCCA 1212
Qy 1141 CCCAGTGAGGACCTCCAGGCGCTGTGGAGTGTGGACATAGTGGGCATGAGTGAAGGC 1200
Db 1213 TGGAGTGAAGACCTCTCAGGCGCTGGGGCAGTGTGGACCTGTGACTATGAGTGAAGCT 1272
Qy 1201 TCAGAAAGCATCTCTCTGCTCATCTGTGCTTGGCTCTGAAAGCCCGCCAGAGGAGCGCTCT 1260
Db 1273 TCAGAAACATCTTCTCTGCTGCTGAGCTTGGCTCTCAAGGCCCGGAGGAGGAGCGCTCA 1332
Qy 1261 GCTGCCAGCTTTGAGTACACTATCTCTGGAGCCCGGAGCTCCCGAGCTCTTGGCTCCATGGACA 1320

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☒ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.